УДК 53(075.8)
Б93
ББК 22.3

БУТИКОВ Е. И., КОНДРАТЬЕВ А. С., УЗДИН В. М. **ФИЗИКА**: Учеб. пособие: В 3 кн. Кн.3. **Строение и свойства вещества.** — М.: ФИЗМАТЛИТ, 2004 — 336 с. — ISBN 5-9221-0109-9 (Кн. 3).

Учебник принципиально нового типа. Последовательность изложения соответствует логической структуре физики как науки и отражает современные тенденции ее преподавания. Материал разделен на обязательный и дополнительный, что позволяет строить процесс обучения с учетом индивидуальных способностей учащихся, включая организацию их самостоятельной работы. Задачи служат как для получения новых знаний, так и для развития навыков исследовательской деятельности.

Для учащихся школ, гимназий, лицеев с углубленным изучением физико-математических дисциплин, а также для подготовки к конкурсным экзаменам в вузы.

Ил. 126.

Рецензенты: профессор *В.И. Николаев*, преподаватель *М.Л. Шифман*

Учебное издание

БУТИКОВ Евгений Иванович
КОНДРАТЬЕВ Александр Сергеевич
УЗДИН Валерий Моисеевич

ФИЗИКА
Книга 3. **Строение и свойства вещества**

Редактор *Д.А. Миртова*
Операторы набора и корректоры: *Л.Г. Быканова, А.К. Розанов*
Оформление обложки *А.Ю. Алехиной*

ЛР № 071930 от 06.07.99.
Подписано в печать 16.02.01. Формат 60×90/16. Бумага офсетная. Печать офсетная
Усл. печ. л. 21,0. Уч.-изд. л. 23,1. Заказ №

Издательская фирма «Физико-математическая литература»
МАИК «Наука/Интерпериодика»
117997, Москва, Профсоюзная, 90
E-mail: fizmat@maik.ru, http://www.fml.ru

Отпечатано с готовых диапозитивов
в ФГУП «Ивановская областная типография»
153008, г. Иваново, ул. Типографская, 6.
E-mail: 091-018@adminet.ivanovo.ru

ISBN 5-9221-0110-2
ISBN 5-9221-0109-9 (Кн. 3)

ОГЛАВЛЕНИЕ

I. ТЕОРИЯ ОТНОСИТЕЛЬНОСТИ

§ 1. Принцип относительности

Теория относительности — это физическая теория пространства и времени, т. е. теория пространственно-временных закономерностей, справедливых для любых физических процессов. Согласно общей теории относительности (называемой также релятивистской теорией тяготения) свойства пространства-времени зависят от действующих в рассматриваемой области полей тяготения. Частная, или специальная, теория относительности описывает свойства пространства-времени в условиях, когда влиянием тяготения на эти свойства можно пренебречь.

Описываемые теорией относительности явления — их называют *релятивистскими* — обнаруживают себя при скоростях движения тел, близких к скорости света в вакууме $c = 3 \cdot 10^8$ м/с. Скорость c — это *предельная скорость* распространения любых взаимодействий и сигналов из одной точки пространства в другую. Понятие предельной скорости требует глубокого изменения обычных (классических) пространственно-временных представлений, основанных на повседневном опыте, ограниченном наблюдениями сравнительно медленных ($v \ll c$) движений.

Все явления в замкнутой физической системе будут происходить точно так же, если всю систему перенести в другое место или как целое повернуть на некоторый угол. В этом проявляются *свойства симметрии* законов природы, отражающие *однородность* пространства (т. е. равноправие всех точек) и его *изотропность* (равноправие всех направлений). Неизменность физических законов с течением времени отражает *однородность времени*.

Наряду с такой инвариантностью (неизменностью) законов природы по отношению к параллельным переносам и поворотам в пространстве и сдвигу во времени на опыте установлена также инвариантность законов физики относительно *преобразований движения*, т. е. перехода от одной инерциальной системы отсчета к другой: все явления в замкнутой физической системе протекают одинаково независимо от того, покоится она в некоторой инерциальной системе отсчета или движется как целое с постоянной скоростью. Это утверждение об эквивалентности (равноправии) всех инерциальных систем отсчета составляет содержание *принципа относительности*.

Рассмотрим лежащие в основе специальной теории относительности физические предпосылки более подробно. Проделав какой-нибудь эксперимент и повторив его при точно таких же условиях в другом месте и в другое время, мы получим тот же самый результат. Этот замечательный факт — воспроизводимость лабораторных опытов — находит свое выражение в независимости физических законов от таких обстоятельств, как положение в пространстве и выбор момента времени. Независимость явлений в замкнутой физической системе от места и момента времени является следствием однородности пространства и времени.

Равноправие инерциальных систем отсчета. Опыт показывает, что, наряду с такой независимостью, существует определенная независимость физических явлений от состояния движения, которая заключается в равноправии всех инерциальных систем отсчета. Равномерное и прямолинейное движение замкнутой системы как целого не влияет на ход процессов, происходящих внутри системы. Это утверждение, впервые высказанное Галилеем для механических явлений, подтверждается всей совокупностью наших знаний о природе и называется принципом относительности.

Многие физические законы формулируются при помощи дифференциальных уравнений. Вид этих уравнений не зависит от начального состояния системы. Таковы, в частности, уравнения механики, которые математически выражают второй закон Ньютона. Согласно принципу относительности математическая форма таких законов должна быть одинакова во всех инерциальных системах отсчета. Другими словами, уравнения движения должны быть *инвариантны* относительно перехода от одной инерциальной системы отсчета к другой.

Понятие события. Рассмотрим описание некоторого явления в двух инерциальных системах отсчета K и K'. Система K' движется относительно K с постоянной скоростью v. Условимся направление одноименных осей в K и K' выбирать одинаковым, а оси x и x' направим вдоль вектора v (рис. 1). Пусть начало отсчета времени $t = 0$ выбрано в тот момент, когда точки O и O' совпадали.

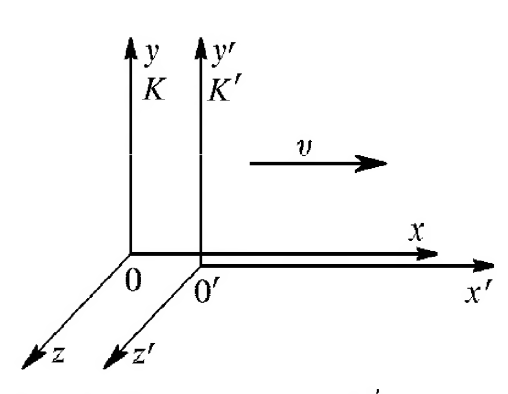


Рис. 1. Система отсчета K' движется относительно K со скоростью v вдоль оси x

Положение некоторой материальной точки определяется координатами и временем x, y, z, t в системе K и координатами и временем x', y', z', t' в другой системе K'. Совокупность трех пространственных координат и времени будем называть *событием*. Таким образом, событие, происходящее с некоторой материальной частицей, определяется местом, где оно произошло, и временем, когда оно произошло.

Классические представления о времени и пространстве. Как связаны между собой координаты и время одного и того же события, если его рассматривать в системах отсчета K и K'? В нерелятивистской физике принималось как очевидный факт существование единого мирового времени t, одинакового во всех системах отсчета: $t = t'$. В действительности возможность измерять время во всех системах отсчета по одним и тем же часам связана с предположением о существовании сигналов, распространяющихся с бесконечно большой скоростью.

Таким образом, согласно классическим представлениям, если два события происходят одновременно в некоторой системе отсчета, то они являются одновременными и в любой другой системе. Точно так же промежуток времени между двумя событиями, в силу абсолютного характера времени, должен быть одинаковым во всех системах отсчета. Предполагалось также, а вернее, считалось очевидным, что длина твердого стержня или вообще расстояние между двумя точками, измеренное в некоторый момент времени, одинаковы во всех системах отсчета.

Преобразования Галилея. Из этих предположений однозначно вытекает общий вид преобразования, связывающего координаты и время некоторого события x,y,z,t в системе K с координатами и временем этого же события x',y',z',t' в системе K'. В самом деле, сравнивая координаты одной и той же частицы в системах отсчета K и K', немедленно получаем

$$x = x' + vt, \qquad y = y', \qquad z = z', \qquad t = t'. \tag{1}$$

Эти формулы носят название *преобразований Галилея.*

Из преобразований Галилея можно сразу получить классический *закон преобразования скорости* частицы при переходе от одной системы отсчета к другой. Пусть $\mathbf{u} = d\mathbf{r}/dt$ — скорость некоторой частицы в K, а $\mathbf{u}' = d\mathbf{r}'/dt'$ — скорость той же частицы в K'. Поскольку $dt = dt'$, из преобразований Галилея получаем

$$u_x = u'_x + v, \qquad u_y = u'_y, \qquad u_z = u'_z. \tag{2}$$

Таким образом, преобразование скорости частицы при переходе от K к K' сводится просто к векторному сложению относительной и переносной скоростей, т. е. к сложению векторов $\mathbf{u}'$ и $\mathbf{v}$.

Уравнение движения классической механики $m\mathbf{a} = \mathbf{F}$ не меняет своего вида при переходе от одной инерциальной системы к другой, т. е. оно инвариантно относительно преобразований Галилея. Другими словами, преобразования Галилея удовлетворяют принципу относительности в отношении законов механики.

Принцип относительности и электродинамика. А как обстоит дело в электродинамике? Что говорит опыт о распространении прин-

ципа относительности на электромагнитные явления? Протекают ли электромагнитные и оптические процессы — взаимодействие зарядов и токов, распространение света — одинаково во всех инерциальных системах отсчета или равномерное прямолинейное движение лаборатории, не оказывая влияния на механические явления, сказывается на электромагнитных явлениях? Вся совокупность экспериментальных данных говорит о том, что принцип относительности распространяется на все явления: как механические, так и электромагнитные и оптические процессы протекают одинаково во всех инерциальных системах отсчета.

Исторически наиболее важные опыты, подтверждающие универсальный характер принципа относительности, — это электродинамический опыт Троутона и Нобля с заряженным конденсатором, подвешенным на упругой нити, и оптический опыт Майкельсона и Морли с интерферометром специальной конструкции. В этих опытах, поставленных специально для обнаружения влияния движения связанной с Землей лаборатории на взаимодействие зарядов и распространение света, был получен отрицательный результат: никакого влияния обнаружено не было.

Однако уравнения электродинамики при переходе от одной инерциальной системы к другой, в отличие от уравнений динамики Ньютона, не являются инвариантными относительно преобразований Галилея. Простые соображения показывают, что преобразования Галилея не удовлетворяют принципу относительности в отношении законов электродинамики и оптики. В самом деле, согласно уравнениям Максвелла, скорость распространения электромагнитных волн, в частности света, в вакууме одинакова по всем направлениям и равна $c = 3 \cdot 10^{10}$ см/с. Но, с другой стороны, в соответствии с классическим законом преобразования скорости, вытекающим из преобразований Галилея, скорость света может быть по всем направлениям равна c только в одной инерциальной системе отсчета. Например, если скорость света равна c в системе K, то в K' свет должен распространяться в положительном направлении оси x' со скоростью $c - v$, а в отрицательном — со скоростью $c + v$. Отсюда можно сделать вывод, что уравнения электродинамики не инвариантны относительно преобразований Галилея.

Таким образом, между электродинамикой и классической механикой имеют место определенные противоречия. Опытные данные свидетельствуют о том, что принцип относительности распространяется на все явления, как механические, так и электродинамические и оптические. В то же время преобразования Галилея удовлетворяют принципу относительности в отношении законов механики и не удовлетворяют в отношении законов электродинамики и оптики.

Отказ от классических представлений. На рубеже XIX и XX веков физика переживала глубокий кризис, единственно правильный рево-

люционный выход из которого был независимо найден такими гигантами, как А. Пуанкаре и А. Эйнштейн, ценой отказа от классических представлений о пространстве и времени и от основанных на них преобразований Галилея. И Эйнштейн, и Пуанкаре опирались на работы Г. А. Лоренца, весьма близко подошедшего к решению проблемы, но не сумевшего сделать последний решительный шаг.

Отказ от преобразований Галилея и введение вместо них новых преобразований — преобразований Лоренца, оставляющих инвариантными при переходе от одной инерциальной системы отсчета к другой уравнения электродинамики, а не уравнения механики, требует пересмотра и уточнения законов классической механики. Решающим шагом на этом пути оказался критический подход к используемому в классической физике понятию абсолютного времени.

Классические представления, почерпнутые из повседневного опыта и кажущиеся наглядными и очевидными, в действительности оказались несостоятельными. Многие понятия и величины, которые в нерелятивистской физике считались абсолютными, т. е. не зависящими от системы отсчета, теория относительности перевела в ранг относительных. Например, считавшееся абсолютным понятие одновременности двух событий в действительности является относительным: два удаленных события, происходящие одновременно в некоторой системе отсчета, не являются одновременными в другой системе, движущейся относительно первой. Промежуток времени между событиями, расстояние между точками в пространстве — эти величины также являются относительными.

Все физические явления происходят в пространстве и во времени; поэтому неудивительно, что внесенное теорией относительности уточнение некоторых основных понятий, в особенности воззрений на пространственные и временные измерения, затронуло в конечном счете всю физику.

Постулаты теории относительности. Теория относительности основана на двух принципах, или постулатах:

1) принцип относительности;

2) принцип существования предельной скорости распространения взаимодействий.

Эти принципы содержат настолько сильные и общие утверждения, что едва ли возможно говорить о каких-либо «решающих» опытах, доказывающих их справедливость. Убеждение в справедливости этих принципов зиждется на бесчисленных опытных проверках следствий теории относительности, которая основана на этих принципах. Сюда относится вся совокупность экспериментальных данных, полученных при изучении движения быстрых частиц в приборах и ускорителях, атомных и ядерных процессов и т. п.

Принцип относительности, как уже отмечалось, есть утверждение об *эквивалентности всех инерциальных систем отсчета*. Равноправие всех инерциальных систем отсчета распространяется на все явле-

ния, на всю физику. Распространение принципа относительности на электромагнитные и оптические явления приводит к выводу о том, что скорость света (электромагнитных волн) в пустоте во всех инерциальных системах отсчета одинакова. Отсюда сразу видна необходимость пересмотра классических представлений о пространстве и времени, так как основанный на них классический закон преобразования скорости находится в противоречии с неизменностью скорости света.

Второй постулат утверждает, что любые взаимодействия между телами распространяются в пустоте с *универсальной конечной скоростью*, не зависящей от движения тел и равной скорости света в вакууме $c = 3 \cdot 10^8$ м/с. Эта скорость определяет тот промежуток времени, после которого до тела может дойти первый сигнал, дающий знать об изменении, происшедшем с другим телом. Значение этого второго постулата связано с тем, что в определении понятий, относящихся к пространству и времени, фундаментальную роль играет передача сигналов с предельной скоростью.

Передача сигналов в принципе возможна не только при помощи электромагнитных волн (света), но и при помощи волн другой природы. Мыслимо, хотя практически и пока неосуществимо, использование гравитационных волн. Не исключено открытие каких-либо новых полей, способных передавать сигналы. Можно, наконец, представить себе передачу сигналов при помощи предельно быстрых частиц. Однако принцип существования универсальной предельной скорости распространения взаимодействий утверждает существование общего предела для скорости передачи каких-либо действий и сигналов и придает скорости света в вакууме универсальное значение, не связанное с физической природой взаимодействия, а отражающее некоторое объективное свойство пространства и времени. Очевидно, что второй постулат утверждает в то же время, что невозможно движение тел со скоростью, превышающей предельную универсальную скорость c.

Второй постулат и законы механики. Отметим, что второй постулат находится в противоречии с принятым в классической механике способом описания взаимодействия материальных частиц, включающим в себя предположение о мгновенности распространения взаимодействий. В самом деле, силы, действующие на каждую из частиц со стороны остальных, считаются в классической механике зависящими от положения частиц в этот же момент времени. Изменение положения какой-либо из частиц мгновенно отражается на остальных. Поэтому второй постулат неизбежно требует уточнения законов механики.

Механика теории относительности в предельном случае, когда скорости движущихся тел малы по сравнению со скоростью света c, переходит в классическую механику, основанную на мгновенности распространения взаимодействий. Только большой величиной скорости распространения взаимодействий объясняется тот факт, что

для макроскопических тел в большинстве случаев достаточно точной оказывается классическая механика. В большинстве случаев скорости, с которыми приходится иметь дело, очень малы по сравнению со скоростью c. Поэтому в то время, когда была создана теория относительности, ее экспериментальное подтверждение можно было найти лишь в исключительно тонких оптических и электродинамических опытах. В настоящее время в больших ускорителях заряженные частицы нередко разгоняются до скоростей, составляющих 99% и более от скорости света. Для расчета траекторий столь быстрых частиц пользоваться механикой Ньютона уже нельзя. В этом смысле можно сказать, что теория относительности в наши дни стала инженерной наукой.

- В чем заключается физическое содержание принципа относительности?
- Что в физике понимают под событием?
- Покажите, что классические представления о пространстве и времени приводят к преобразованиям Галилея.
- Почему преобразования Галилея несовместимы с представлением о существовании конечной предельной скорости распространения взаимодействий?
- Какие постулаты лежат в основе теории относительности?
- Почему предположение о конечности скорости распространения взаимодействий требует уточнения законов классической механики?

§ 2. Релятивистская кинематика

Постулаты теории относительности требуют внесения изменений в основные физические понятия, относящиеся к пространству и времени. Прежде всего необходим анализ основных измерительных операций, определяющих пространственно-временные соотношения между событиями. Главное, что внесла теория относительности в постановку вопроса об измерительных операциях, состоит в том, что любое физическое понятие (например, измерение промежутков времени и расстояний) нуждается в определении.

Процедуры измерений. Измерение времени может быть, в принципе, произведено при помощи любого периодического процесса. Наибольшей точностью в настоящее время обладают часы, основанные на использовании собственных колебаний молекул аммиака (молекулярные часы) или атомов цезия (атомные часы). Основанное на колебаниях атомов измерение времени особенно удобно тем, что здесь природа предоставила нам, в силу тождественности атомов одного и того же изотопа, набор совершенно идентичных часов. Выяснять, идут ли выбранные в качестве эталона часы равномерно, — бессмысленно: это по определению так.

Одновременность событий и синхронизация часов. Что значит измерить промежуток времени между событиями? Это значит сравнить между собой показания выбранных в качестве эталона часов в моменты наступления этих событий. А для этого нужно установить одновременность рассматриваемого события с другим событием — прохождением стрелки часов через определенное деление. Таким образом, все наши суждения, в которых время играет какую-либо роль, всегда являются суждениями об *одновременных событиях*.

Понятие одновременности событий, происходящих в одном и том же месте, «рядом», не нуждается в определении. Поэтому для измерения промежутка времени между двумя событиями, происходящими в одном месте, достаточно иметь в этом же месте часы. Но как быть с «удаленными» событиями, происходящими в разных местах? Для измерения промежутка времени между такими событиями нужно иметь в тех точках, где они происходят, синхронно идущие идентичные часы.

Но как узнать, что находящиеся в различных точках A и B часы идут синхронно, или, что то же самое, как узнать, что два определенных события в точках A и B происходят одновременно? *Узнать* это нельзя, нужно сначала дать *определение*, что такое одновременность для пространственно удаленных точек. Без такого определения невозможно сравнивать по времени события, происходящие в различных точках.

Эйнштейновское определение одновременности основано на независимости скорости сигнала, распространяющегося с предельной скоростью, от направления. Пусть из точки A в момент времени t_1 по часам A отправляется сигнал (рис. 2). Пусть время прихода сиг-

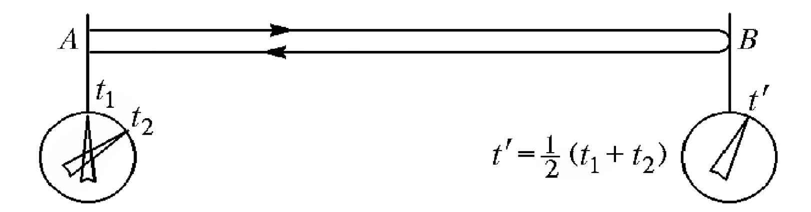


Рис. 2. К определению понятия одновременности событий

нала в B и отражения назад на часах B есть t'. Наконец, пусть отраженный сигнал приходит в A в момент t_2 по часам A. Тогда по определению часы в A и B идут синхронно, если $t' = (t_1 + t_2)/2$.

Как уже отмечалось, в нерелятивистской физике принималось как нечто само собой разумеющееся существование единого мирового времени, не зависящего от системы отсчета, и поэтому неявно допускалось, что понятие одновременности в разных точках пространства не нуждается в определении, а любой способ синхронизации часов (путем световых сигналов или путем перевозки хронометров) должен

дать одно и то же. Мы увидим, что на самом деле это не так. Если часы в B синхронизированы с часами в A с помощью световых сигналов, и хронометр C, сверенный с часами в точке A, перевозится затем в точку B, то его показания в точке B не совпадут с показаниями находящихся там часов, а будут зависеть от скорости перевозки. Совпадение будет лишь при бесконечно малой скорости перевозки хронометра.

Измерение расстояний. Операцию измерения расстояний с точки зрения постулатов теории относительности наиболее разумно определить на основе «радиолокационного» способа: из некоторого пункта посылаются световые или радиосигналы, которые после отражения от наблюдаемого предмета возвращаются в точку отправления. При этом измеряется время прохождения сигнала туда и обратно по часам, связанным с радиолокатором. Расстояние до предмета получают, умножая одинаковую по всем направлениям скорость сигнала на половину времени прохождения: $l = c(t_2 - t_1)/2$.

В принципиальном отношении этот способ важен потому, что в нем измерение расстояний сводится к измерению времени и отпадает необходимость в отдельном эталоне длины.

Радиолокационный способ измерения расстояний не является единственно возможным: можно было бы по определению в качестве измерительной процедуры принять способ непосредственного наложения жестких масштабов (линеек) или способ триангуляции. Однако эти способы менее удовлетворительны, так как существенно опираются на свойства твердых тел. Но абсолютно твердых тел в природе не существует, реальные физические тела лишь приближенно могут рассматриваться как твердые и обладающие неизменными геометрическими размерами.

Относительность одновременности событий. До сих пор наши рассуждения относились к какой-либо одной инерциальной системе отсчета. Будем теперь рассматривать события, промежутки времени и расстояния с точки зрения разных систем отсчета.

В нерелятивистской физике время является абсолютным: для всех систем отсчета вводится одно и то же время. Это значит, что если два события происходят одновременно для какого-нибудь наблюдателя, то они являются одновременными и для любого другого: понятие одновременности является абсолютным, не зависящим от системы отсчета. Однако утверждение об абсолютном характере одновременности основано на предположении о существовании сигналов, распространяющихся мгновенно, с бесконечно большой скоростью.

Покажем, что второй постулат теории относительности, утверждающий существование предельной конечной скорости сигналов, выражает относительный характер одновременности. Утверждение, что два пространственно удаленных события происходят одновременно, приобретает смысл только тогда, когда указано, к какой системе отсчета это утверждение относится.

Рассмотрим опять две инерциальные системы отсчета K и K', причем K' движется относительно K в положительном направлении оси x (рис. 3). Пусть из некоторой точки А на оси x' отправляются сигналы во взаимно противоположных направлениях. Рассмотрим приход сигналов в точки B и C системы K', равноудаленные от A. Очевидно, что сигналы достигнут B и C одновременно по часам системы K'.

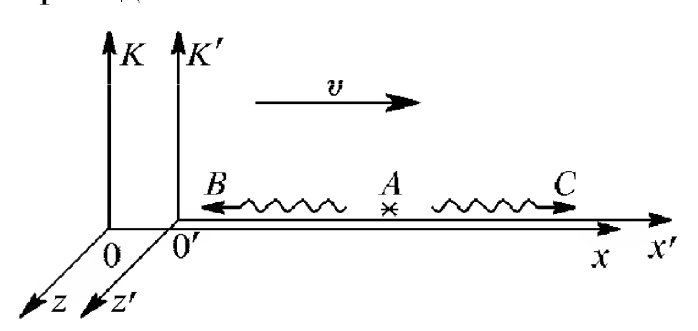


Рис. 3. Относительный характер одновременности событий

Легко видеть, однако, что эти же два события — приход сигналов в B и C, — одновременные в K', отнюдь не будут одновременными для наблюдателя в системе K. В самом деле, согласно принципу относительности, скорость сигналов в системе K также не зависит от направления, но точка B относительно K движется направо, навстречу посланному в нее сигналу, а точка C — по направлению от посланного в нее сигнала. Поэтому, с точки зрения наблюдателя в K, сигналу, распространяющемуся с конечной скоростью, приходится на пути в В преодолевать меньшее расстояние, чем на пути в C. Следовательно, в системе K сигнал в точку В придет раньше, чем в C, и, значит, понятие одновременности событий является относительным.

Относительность промежутков времени. Покажем теперь исходя из основных постулатов теории относительности относительный характер промежутков времени между событиями. Пусть два события в некоторой системе отсчета, скажем в K', происходят в одной и той же точке и промежуток времени между ними равен τ_0 по часам системы K'. Этот промежуток времени называется *собственным временем*. Каким будет промежуток времени между этими же событиями, если его измерить по часам системы K, относительно которой K' движется со скоростью v?

Для ответа на этот вопрос рассмотрим мысленный опыт со «световыми часами», устроенными следующим образом (рис. 4). На концах стержня длины l закреплены два параллельных зеркала. Между зеркалами движется короткий световой импульс. Пусть этот прибор неподвижен в системе K' и его стержень расположен перпендикулярно скорости системы отсчета K' относительно K.

Рассмотрим один цикл таких часов, т. е. выход светового импульса от нижнего зеркала и его возвращение после отражения от верхнего зеркала, с точки зрения каждой из систем отсчета. В системе K' оба рассматриваемых события происходят в одной и той же точке и промежуток времени между ними (собственное время) равен $\tau_0 = 2l/c$.

С точки зрения системы K часы находятся в движении и световой импульс движется между зеркалами зигзагообразно (рис. 4). Свет при этом проходит за один цикл больший путь, и, следователь-

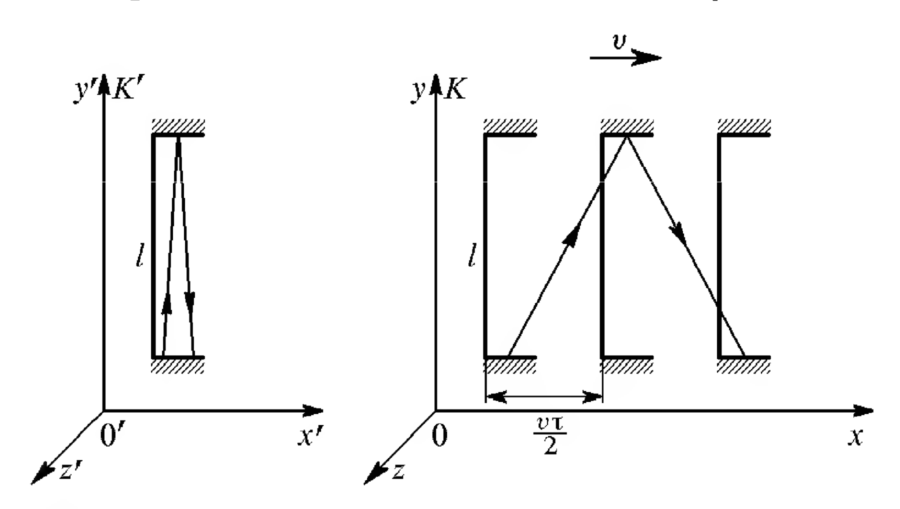


Рис. 4. Свет проходит от одного зеркала до другого и обратно за разное время в различных системах отсчета

но, промежуток времени τ между этими же событиями, измеренный в системе отсчета K, больше, чем в K': $\tau > \tau_0$. В этих рассуждениях мы опираемся на то, что, согласно второму постулату, скорость света c одинакова в K и K'.

Найдем связь τ и τ_0. Как видно из рис. 4, пройденный светом за один цикл путь равен $2\sqrt{l^2 + (v\tau/2)^2}$, и для определения τ можно написать уравнение

$$c\tau = 2\sqrt{l^2 + (v\tau/2)^2},$$

откуда

$$\tau = \frac{2l}{c}\frac{1}{\sqrt{1 - v^2/c^2}}.$$

Но, как мы видели выше, $2l/c$ равно промежутку времени τ_0 между этими событиями в K'. Поэтому

$$\tau = \frac{\tau_0}{\sqrt{1 - v^2/c^2}}. \tag{1}$$

Таким образом, промежуток времени между двумя событиями зависит от системы отсчета, т. е. является относительным. Так как при любой отличной от нуля скорости $\tau > \tau_0$, то собственное время меньше, чем промежуток времени между этими же событиями, измеренный в любой другой системе отсчета. Этот эффект называют *релятивистским замедлением* или «растяжением» времени. С точки зрения наблюдателя K идентичные по устройству движущиеся часы (т. е. часы в K') идут медленнее, чем его собственные.

Подчеркнем, что замедление времени является следствием инвариантности скорости света. Оно, разумеется, не зависит от конкретной конструкции часов.

Рассмотренный релятивистский эффект замедления времени является взаимным, как того требует принцип относительности, т. е. постулат о равноправии инерциальных систем отсчета K и K': с точки зрения наблюдателя в системе K' медленнее идут часы, связанные с системой K.

Отметим, что в приведенных рассуждениях длину стержня l в направлении, перпендикулярном относительной скорости систем отсчета K и K', мы считали одинаковой в обеих системах отсчета. Если предположить, что это не так, то можно сразу прийти к противоречию с равноправием систем K и K'. В самом деле, рассмотрим следующий мысленный опыт. Расположим вдоль оси y' системы K' жесткий стержень, длина которого в этой системе равна l, и вдоль оси y системы K расположим точно такой же стержень, т. е. длина этого стержня равна l для наблюдателя в K. В некоторый момент эти стержни оказываются рядом, и представляется возможность сравнить их непосредственно — конец одного из стержней может сделать метку на другом стержне. Совпадет ли эта метка с концом стержня? Принцип относительности дает положительный ответ на этот вопрос: метка совпадает с концом стержня, т. е. длина стержня в направлении, перпендикулярном относительной скорости систем отсчета K и K', одинакова в обеих системах. Если бы совпадения не было, то один из стержней оказался бы длиннее другого с точки зрения обеих систем отсчета, что противоречит принципу относительности.

Относительность пространственных расстояний. Покажем теперь, что длина твердого стержня, расположенного вдоль направления относительной скорости систем отсчета K и K' (рис. 5), будет различной в этих системах. Пусть стержень покоится в системе отсчета K'. Его длину, измеренную в этой системе отсчета, называют собственной длиной. Обозначим ее через l_0, а длину в системе K, относительно которой стержень движется со скоростью v, через l. Найдем связь между l и l_0.

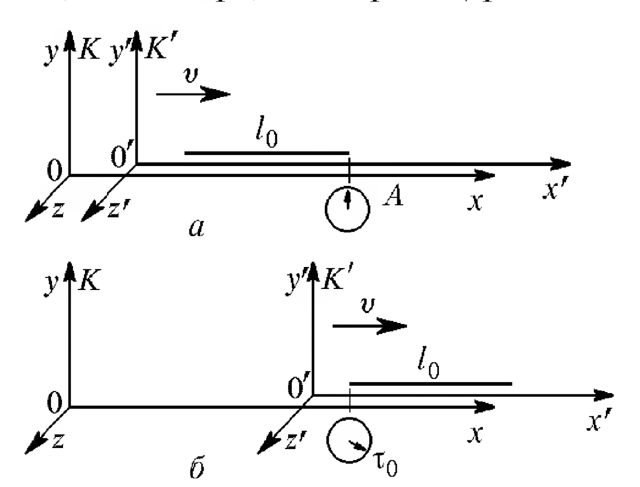


Рис. 5. Длина твердого стержня различна в разных системах отсчета

Для этого рассмотрим два события: *а*) прохождение начала стержня мимо точки A на оси x системы K и *б*) прохождение конца стержня мимо этой же точки. В системе K

эти события происходят в одной точке, и промежуток времени между ними в системе K является собственным временем τ_0. Так как стержень движется относительно K со скоростью v, то можно написать $l = v\tau_0$. Но с точки зрения наблюдателя в системе K' точка А движется вдоль неподвижного стержня налево с такой же скоростью, поэтому $l_0 = v\tau$, где τ есть промежуток времени между событиями *а* и *б*, измеренный по часам в K'. Так как $\tau = \tau_0/\sqrt{1 - v^2/c^2}$, то, комбинируя соотношения $l = v\tau_0$ и $l_0 = v\tau$, находим

$$l = l_0\sqrt{1 - \frac{v^2}{c^2}}. \quad (2)$$

Мы приходим к выводу, что длина стержня зависит от системы отсчета, в которой она измеряется, т. е. является относительной. При любой отличной от нуля скорости $l < l_0$, т. е. длина стержня является наибольшей в той системе отсчета, в которой стержень покоится. Движущиеся относительно наблюдателя тела сокращаются в направлении своего движения. Этот релятивистский эффект носит название *лоренцева сокращения.*

Лоренцево сокращение движущегося стержня отражает относительный характер расстояния между точками в теории относительности (т. е. зависимость расстояния от системы отсчета) и не связано с какими-либо процессами или явлениями в самом стержне. Оно тем не менее представляет собой вполне реальный эффект, столь же реальный, как, например, зависимость скорости тела от выбора системы отсчета.

В полном соответствии с принципом относительности эффект сокращения длины стержня является взаимным: если такой же стержень покоится в системе отсчета K, то его длина в этой системе отсчета равна l_0, а в системе K' длина будет меньше в соответствии с приведенной формулой.

Классический предел. Как видно из полученной формулы, эффект сокращения длины зависит от относительной скорости v систем отсчета и становится особенно заметным для скоростей, сравнимых со скоростью света. При $v \to c$ длина стремится к нулю: $l \to 0$.

Зависимость лоренцева сокращения от скорости показана на рис. 6. При малых скоростях ($v \ll c$) $l \approx l_0$ и $\tau \approx \tau_0$, т. е. расстояние между точками и промежуток времени между событиями приобретают практически абсолютный смысл в полном соответствии с классическими представлениями о пространстве и времени, сформировавшимися на основе многовекового опыта наблюдений над сравнительно медленными движениями, происходящими со скоростями, малыми по сравнению со скоростью света.

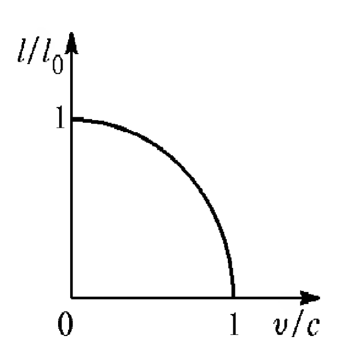


Рис. 6. Зависимость сокращения длины от относительной скорости

- Какие события считаются одновременными? Как производится синхронизация часов?
- В чем заключается относительный характер одновременности событий?
- Что называется собственным временем и собственной длиной?
- Получите выражения для релятивистских эффектов замедления времени и сокращения длины, исходя непосредственно из постулатов теории относительности.
- Получите формулу для лоренцева сокращения длины, рассматривая цикл световых часов, расположенных вдоль относительной скорости двух систем отсчета.

§ 3. Преобразования Лоренца

Полученные в § 2 на основе постулатов теории относительности формулы (1) и (2), связывающие промежутки времени и расстояния между точками в разных системах отсчета, позволяют написать релятивистский закон преобразования координат и времени произвольного события при переходе от одной системы отсчета к другой. Этот закон должен заменить основанные на классических представлениях о пространстве и времени преобразования Галилея.

Рассмотрим, как и в § 1, описание некоторого события А в двух инерциальных системах отсчета K и K'. Пусть координаты и время этого события в системе K есть x, y, z и t, а в системе K' — x', y', z' и t' (рис. 7). Как и прежде, считаем, что при $t = 0$ точки O и O' совпадают. Расстояния в направлении, перпендикулярном относительной скорости v систем отсчета, как уже было показано, одинаковы в K и K', поэтому $y = y'$ и $z = z'$.

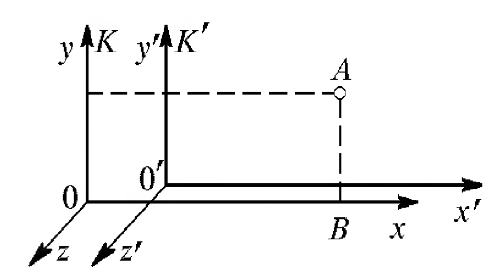


Рис. 7. Координаты одного и того же события A в двух системах отсчета

Координата x есть собственная длина l_0 отрезка OB, неподвижного в K-системе. Длина l этого же отрезка в K'-системе, где измерение производится в момент времени t', есть $x' + vt'$. Учитывая соотношение (2) предыдущего параграфа между собственной длиной некоторого отрезка l_0 и длиной l этого же отрезка в системе отсчета, относительно которой он движется со скоростью v:

$$l = l_0\sqrt{1 - \frac{v^2}{c^2}},$$

мы можем написать

$$x' + vt' = x\sqrt{1 - \frac{v^2}{c^2}},$$

откуда

$$x = \frac{x' + vt'}{\sqrt{1 - v^2/c^2}}. \qquad (1)$$

Но можно рассуждать и иначе. Координата x' есть собственная длина отрезка $O'B$, неподвижного в K'-системе. Длина этого же отрезка в K, измеряемая в момент времени t по часам K, равна $x - vt$. Снова учитывая соотношение между длиной одного и того же отрезка в двух системах отсчета, можем написать

$$x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}}. \qquad (2)$$

Формулы (1) и (2) представляют собой искомый закон преобразования координаты x при переходе от K' к K и обратно. Взятые вместе, они позволяют найти связь между временем t и t' одного и того же события в обеих системах отсчета. Исключая из (1) и (2) сначала x', а затем x, найдем

$$t = \frac{t' + \frac{v}{c^2}x'}{\sqrt{1 - v^2/c^2}}, \qquad t' = \frac{t - \frac{v}{c^2}x}{\sqrt{1 - v^2/c^2}}. \qquad (3)$$

Преобразования Лоренца. Интервал между событиями. Таким образом, релятивистские формулы преобразования координат некоторого события при переходе от одной инерциальной системы отсчета к другой имеют вид

$$x = \frac{x' + vt'}{\sqrt{1 - v^2/c^2}}, \qquad y = y', \qquad z = z', \qquad t = \frac{t' + \frac{v}{c^2}x'}{\sqrt{1 - v^2/c^2}}. \qquad (4)$$

Эти формулы называют *преобразованиями Лоренца*. Они заменяют преобразования Галилея, справедливые лишь в предельном случае малых по сравнению со скоростью света относительных скоростей. При $v \ll c$ преобразования Лоренца (4) переходят в преобразования Галилея. Это означает, что теория относительности не отвергает полностью классические представления о пространстве и времени, а включает их в себя как предельный случай, справедливый для медленных движений. Теория относительности не отвергает классическую физику, а определяет границы ее применимости.

Преобразования Лоренца выражают относительный характер промежутков времени между событиями и расстояний между точками в пространстве. Однако наиболее характерной чертой теории относительности является не утверждение относительного характера пространства и времени, а установление абсолютных, не зависящих от выбора систем отсчета законов природы.

Задача нахождения абсолютного выражения законов природы тесно связана с отысканием абсолютных, инвариантных величин. Одна из таких величин упоминается уже в основных постулатах — это максимальная скорость распространения взаимодействий, равная скорости света в вакууме c. Другой важной инвариантной вели-

чиной является *пространственно-временной интервал* между событиями, определяемый следующим соотношением:

$$S_{12} = \sqrt{c^2 t_{12}^2 - l_{12}^2}, \tag{5}$$

где t_{12} — промежуток времени между событиями в некоторой системе отсчета, а l_{12} — выраженное в той же системе расстояние между точками, в которых происходят рассматриваемые события. В частности, если одному из происходящих событий соответствуют координаты $x_1 = y_1 = z_1 = 0$ и момент времени $t_1 = 0$, а второму — координаты x, y, z и момент t, то интервал между ними

$$S = \sqrt{c^2 t^2 - x^2 - y^2 - z^2}. \tag{6}$$

Пусть, например, первое событие представляет собой вспышку света, происходящую в начале координат при $t = 0$, а второе — приход фронта этой световой волны в точку с координатами x, y, z в момент времени t. Тогда $x^2 + y^2 + z^2 = c^2 t^2$ и интервал для такой пары событий $S = 0$. Координаты и время второго события в другой системе отсчета K' будут другими, но и для них в силу инвариантности скорости света будет выполняться такое же соотношение $x'^2 + y'^2 + z'^2 = c^2 t'^2$ и следовательно, $S' = 0$. Таким образом, если два события связаны между собой световым сигналом, то интервал между ними равен нулю во всех инерциальных системах отсчета. Этот результат является математическим выражением абсолютного характера скорости света.

Инвариантность интервала. Для любой другой пары событий, не связанных световым сигналом, интервал отличен от нуля, но величина его во всех инерциальных системах отсчета одинакова:

$$c^2 t^2 - x^2 - y^2 - z^2 = c^2 t'^2 - x'^2 - y'^2 - z'^2.$$

В этом легко убедиться с помощью преобразований Лоренца (4), подставив в левую часть выражения для x, y, z и t через координаты и время этого же события x', y', z' и t' в другой системе отсчета. Инвариантность интервала означает, что утверждение «два события разделены интервалом S» имеет абсолютный характер, т. е. оно справедливо во всех инерциальных системах отсчета.

Классификация интервалов. Понятие интервала между событиями является обобщением понятий промежутка времени и расстояния между точками. В зависимости от того, какая составляющая — временная или пространственная — преобладает в рассматриваемом интервале, возникает деление интервалов на времениподобные и пространственноподобные. Для времениподобного интервала $c^2 t_{12}^2 > l_{12}^2$, т. е. $S_{12}^2 > 0$. В этом случае всегда можно найти такую систему отсчета K', в которой рассматриваемые события происходят

в одной точке, т. е. $l'_{12} = 0$, и промежуток времени между ними в такой системе отсчета является собственным временем $t'_{12} = \tau_0$:

$$S_{12}^2 = c^2 t_{12}^2 - l_{12}^2 = c^2 t_{12}'^2 - l_{12}'^2 = c^2 \tau_0^2.$$

Таким образом, для событий, разделенных времениподобным интервалом, всегда существует такая система отсчета, в которой этот интервал (с точностью до постоянного множителя c) представляет собой просто промежуток времени τ_0 между этими событиями. Для этих событий понятия «раньше», «позже» имеют абсолютный характер. Очевидно, что между такими событиями может иметь место причинно-следственная связь.

Для событий, разделенных пространственноподобным интервалом, $c^2 t_{12}^2 < l_{12}^2$, т. е. $S_{12}^2 < 0$, а интервал является мнимым числом. В этом случае всегда можно найти такую систему отсчета K', в которой эти события происходят одновременно, т. е. $t'_{12} = 0$:

$$c^2 t_{12}^2 - l_{12}^2 = c^2 t_{12}'^2 - l_{12}'^2 = -l_{12}'^2.$$

Абсолютная величина пространственноподобного интервала представляет собой пространственное расстояние между событиями в той системе отсчета, в которой эти события произошли одновременно. Понятия «одновременно», «раньше», «позже» для таких событий относительны: всегда можно указать такие системы отсчета, в которых первое событие происходит раньше второго, и такие, в которых второе происходит раньше первого. Ясно, что между событиями, для которых теряют абсолютный смысл понятия «раньше» и «позже», не может быть причинно-следственной связи.

Впрочем, невозможность причинной связи между событиями, разделенными пространственноподобным интервалом, для которых $l_{12} > ct_{12}$, непосредственно видна из того, что никакой сигнал, никакое взаимодействие не может распространяться со скоростью, большей c.

Равный нулю интервал между событиями, связанными световым сигналом, называют светоподобным.

Подчеркнем еще раз, что разделение интервалов между событиями на времениподобные и пространственноподобные вследствие инвариантности интервала является абсолютным, т. е. не зависящим от системы отсчета.

Закон преобразования скорости. Преобразования Лоренца для координат и времени события при переходе от одной инерциальной системы отсчета к другой позволяют сразу получить и закон преобразования скорости частицы. Пусть некоторая частица за малый промежуток времени Δt по часам системы отсчета K переместилась из точки с радиусом-вектором $\mathbf{r}$ в точку с радиусом-вектором $\mathbf{r} + \Delta\mathbf{r}$. Тогда по определению скоростью $\mathbf{u}$ этой частицы относительно системы K называется

предел отношения $\Delta\mathbf{r}/\Delta t$ при $\Delta t \to 0$. Скоростью $\mathbf{u}'$ этой же частицы в системе отсчета K' будет предел отношения $\Delta\mathbf{r}'/\Delta t'$, где $\Delta t'$ и $\Delta\mathbf{r}'$ — промежуток времени и изменение радиуса-вектора в K' для той же пары событий, т. е. для конечного и начального положений частицы. Применяя преобразования Лоренца (4) к конечному и начальному положениям частицы и вычитая их почленно, получим

$$\Delta x = \frac{\Delta x' + v\Delta t'}{\sqrt{1 - v^2/c^2}}, \qquad \Delta y = \Delta y', \qquad \Delta z = \Delta z', \qquad \Delta t = \frac{\Delta t' + \dfrac{v}{c^2}\,\Delta x'}{\sqrt{1 - v^2/c^2}}.$$

Разделив почленно первые три равенства на четвертое, находим

$$\frac{\Delta x}{\Delta t} = \frac{\dfrac{\Delta x'}{\Delta t'} + v}{1 + \dfrac{v}{c^2}\dfrac{\Delta x'}{\Delta t'}}, \qquad \frac{\Delta y}{\Delta t} = \frac{\dfrac{\Delta y'}{\Delta t'}\sqrt{1 - \dfrac{v^2}{c^2}}}{1 + \dfrac{v}{c^2}\dfrac{\Delta x'}{\Delta t'}}, \qquad \frac{\Delta z}{\Delta t} = \frac{\dfrac{\Delta z'}{\Delta t'}\sqrt{1 - \dfrac{v^2}{c^2}}}{1 + \dfrac{v}{c^2}\dfrac{\Delta x'}{\Delta t'}}.$$

Переходя в этих формулах к пределу при $\Delta t \to 0$, $\Delta t' \to 0$, получим

$$u_x = \frac{u'_x + v}{1 + \dfrac{v}{c^2}u'_x}, \qquad u_y = \frac{u'_y\sqrt{1 - \dfrac{v^2}{c^2}}}{1 + \dfrac{v}{c^2}\,u'_x}, \qquad u_z = \frac{u'_z\sqrt{1 - \dfrac{v^2}{c^2}}}{1 + \dfrac{v}{c^2}\,u'_x}. \qquad (7)$$

Выражения (7) представляют собой закон преобразования скорости частицы при переходе от одной инерциальной системы отсчета к другой. Отметим, что поперечные к направлению относительной скорости систем отсчета компоненты скорости частицы u_y и u_z, в отличие от поперечных координат y и z, не остаются неизменными. Это связано с тем, что при переходе от одной системы отсчета к другой время преобразуется.

В предельном случае $v \ll c$ релятивистские формулы (7) переходят в формулы классической механики:

$$u_x = u'_x + v, \qquad u_y = u'_y, \qquad u_z = u'_z.$$

Релятивистский закон преобразования скорости согласуется, разумеется, с исходным постулатом об инвариантности скорости света. Рассмотрим, например, в системе отсчета K' световой импульс, распространяющийся вдоль оси x'. Для такого импульса $u'_x = c$, $u'_y = u'_z = 0$. Тогда, согласно (7), для скорости этого же импульса в системе отсчета K получим

$$u_x = \frac{c + v}{1 + v/c} = c, \qquad u_y = 0, \qquad u_z = 0,$$

т. е. импульс u в системе K распространяется вдоль оси x со скоростью c.

Аберрация света. В качестве примера применения релятивистского преобразования скорости рассмотрим точечный источник света, покоящийся в системе K' и равномерно излучающий свет по всем направлениям (рис. 8). Рассмотрим те 50% светового потока, которые источник излучает в переднюю полусферу в системе отсчета K'. С точки зрения наблюдателя в системе K, относительно которой источник движется со скоростью v, излучение уже отнюдь не будет изотропным: эти 50% будут излучаться преимущественно вперед в конус с углом δ (рис. 9), причем $\cos\delta = v/c$.

Рис. 8. Неподвижный источник света излучает равномерно по всем направлениям

В самом деле, в системе K' луч света, ограничивающий рассматриваемый пучок, направлен вдоль оси y' и для него $u'_x = 0$, $u'_y = c$, $u'_z = 0$. Переходя в систему отсчета K, мы для этого же луча получим, согласно формулам (7),

$$u'_x = v, \qquad u'_y = \sqrt{1 - v^2/c^2}, \qquad u'_z = 0,$$

откуда

$$\cos\delta = \frac{u_x}{\sqrt{u_x^2 + u_y^2}} = \frac{v}{c}.$$

При скорости источника v, близкой к скорости света, пучок света сконцентрируется в узкий конус, направленный вперед по движению, с осью, совпадающей с направлением движения источника.

Совершенно аналогично с помощью релятивистского закона преобразования скорости объясняется явление звездной аберрации. Движущийся наблюдатель обнаружит искажение картины звездного неба: для него Вселенная «сжата» в направлении его движения по сравнению с картиной, которую видит в том же направлении неподвижный относительно звезд наблюдатель. Если движущийся наблюдатель будет менять свою скорость, то он обнаружит, что звездное небо «переливается» вокруг него: направления, в которых он видит звезды, будут все время меняться, не образуя постоянных углов друг с другом. Именно в таком положении находится наблюдатсль на Зсмлс, обращающсйся вокруг Солнца. Каждые полгода скорость Земли в ее годичном движении (30 км/с, т. е. $10^{-4}c$) меняет направление на противоположное. В аберрации света звезд обнаруживается, конечно, не сама скорость Земли, а тот факт, что эта

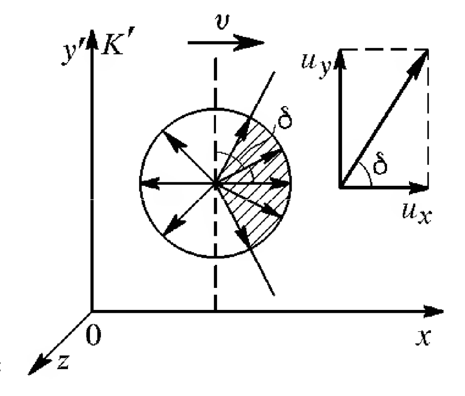


Рис. 9. Для наблюдателя, относительно которого источник света движстся, излучснис нс является изотропным

скорость изменяется: в разное время года положения звезд сдвинуты по-разному.

- Покажите, что в предельном случае малых скоростей $v \ll c$, преобразования Лоренца переходят в преобразования Галилея.
- Получите из формул преобразований Лоренца выражения для релятивистского преобразования промежутков времени между двумя событиями и для лоренцева сокращения длины движущегося стержня.
- Что такое интервал между двумя событиями? Докажите его инвариантность.
- В каких случаях интервал между двумя событиями называется светоподобным? времениподобным? пространственноподобным? Поясните связь этой классификации интервалов с возможностью причинно-следственной связи между событиями.
- Почему при движении системы отсчета K' вдоль оси x системы преобразуются не только x-проекция скорости частицы, но также и y- и z-проекции?
- Как релятивистская кинематика объясняет явление звездной аберрации?

§ 4. Релятивистская динамика

Теория относительности требует пересмотра и уточнения законов механики. Как мы видели, уравнения классической динамики (второй закон Ньютона) удовлетворяют принципу относительности в отношении преобразований Галилея. Но ведь преобразования Галилея должны быть заменены преобразованиями Лоренца! Поэтому уравнения динамики следует изменить так, чтобы они оставались неизменными при переходе от одной инерциальной системы отсчета к другой, согласно преобразованиям Лоренца. При малых скоростях $(v \ll c)$ уравнения релятивистской динамики должны переходить в классические, ибо в этой области их справедливость подтверждается на опыте.

Импульс и энергия. В теории относительности, как и в классической механике, для замкнутой физической системы сохраняются импульс $\mathbf{p}$ и энергия E, однако релятивистские выражения для них отличаются от соответствующих классических:

$$\mathbf{p} = \frac{m_0 \mathbf{v}}{\sqrt{1 - v^2/c^2}}, \qquad E = \frac{m_0 c^2}{\sqrt{1 - v^2/c^2}}. \tag{1}$$

здесь m_0 — масса частицы. Это масса в той системе отсчета, где частица покоится. Часто ее называют *массой покоя* частицы. Она совпадает с массой частицы в нерелятивистской механике.

Можно показать, что выражаемая формулами (1) зависимость импульса и энергии частицы от ее скорости в теории относительности с неизбежностью следует из релятивистского эффекта замедления времени в движущейся системе отсчета. Это будет сделано ниже.

Релятивистские энергия и импульс (1) удовлетворяют уравнениям, аналогичным соответствующим уравнениям классической механики:

$$\frac{d\mathbf{p}}{dt} = \sum \mathbf{F}, \qquad \frac{dE}{dt} = \sum \mathbf{F}\cdot\mathbf{v}. \tag{2}$$

Релятивистская масса. Иногда коэффициент пропорциональности в (1) между скоростью частицы и ее импульсом

$$m = \frac{m_0}{\sqrt{1 - v^2/c^2}} \tag{3}$$

называют *релятивистской массой* частицы. С ее помощью выражения (1) для импульса и энергии частицы можно записать в компактном виде

$$\mathbf{p} = m\mathbf{v}, \qquad E = mc^2. \tag{4}$$

Если релятивистской частице, т. е. частице, движущейся со скоростью, близкой к скорости света, сообщить дополнительную энергию, чтобы увеличить ее импульс, то скорость ее при этом увеличится очень незначительно. Можно сказать, что энергия частицы и ее импульс увеличиваются теперь за счет роста ее релятивистской массы. Этот эффект наблюдается в работе ускорителей заряженных частиц высоких энергий и служит наиболее убедительным экспериментальным подтверждением теории относительности.

Энергия покоя. Самое замечательное в формуле $E = mc^2$ заключается в том, что покоящееся тело обладает энергией: полагая в (1) $v = 0$, получаем

$$E_0 = m_0c^2. \tag{5}$$

Энергию E_0 называют *энергией покоя.*

Кинетическая энергия. Кинетическая энергия частицы $E_к$ в некоторой системе отсчета определяется как разность между ее полной энергией $E = mc^2$ и энергией покоя $E_0 = m_0c^2$. С помощью (1) имеем

$$E_к = mc^2 - m_0c^2 = m_0c^2\left(\frac{1}{\sqrt{1 - v^2/c^2}} - 1\right). \tag{6}$$

Если скорость частицы мала по сравнению со скоростью света, формула (6) переходит в обычное выражение $E_к = m_0v^2/2$ для кинетической энергии частицы в нерелятивистской физике.

Различие между классическим и релятивистским выражениями для кинетической энергии становится особенно существенным, когда скорость частицы приближается к скорости света. При $v \to c$ релятивистская кинетическая энергия (6) неограниченно возрастает: частица, обладающая отличной от нуля массой покоя m_0 и движу-

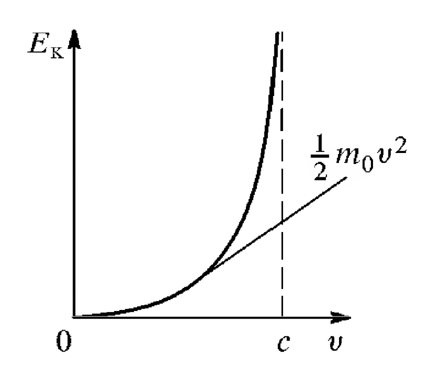


Рис. 10. Зависимость кинетической энергии тела от скорости

щаяся со скоростью света, должна была бы иметь бесконечную кинетическую энергию. Зависимость кинетической энергии от скорости частицы показана на рис. 10.

Пропорциональность массы и энергии. Из формулы (6) следует, что при разгоне тела приращение кинетической энергии сопровождается пропорциональным приращением его релятивистской массы. Вспомним, что важнейшим свойством энергии является ее способность превращаться из одной формы в другую в эквивалентных количествах при различных физических процессах — именно в этом заключается содержание закона сохранения энергии. Поэтому естественно ожидать, что возрастание релятивистской массы тела будет происходить не только при сообщении ему кинетической энергии, но и при любом другом увеличении энергии тела, независимо от конкретного вида энергии. Отсюда можно сделать фундаментальное заключение о том, что полная энергия тела пропорциональна его релятивистской массе независимо от того, из каких конкретных видов энергии она состоит.

Поясним сказанное на следующем простом примере. Рассмотрим неупругое столкновение двух одинаковых тел, движущихся навстречу друг другу с одинаковыми скоростями, так что в результате столкновения образуется одно тело, которое покоится (рис. 11*а*). Пусть

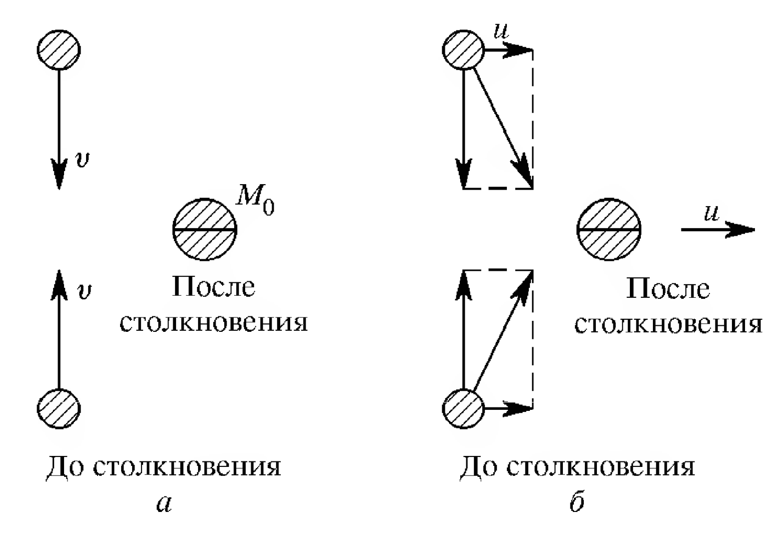


Рис. 11. Неупругое столкновение, наблюдаемое в разных системах отсчета

скорость каждого из тел до столкновения равна v, а масса покоя m_0. Массу покоя образовавшегося тела обозначим через M_0. Теперь рассмотрим это же столкновение с точки зрения наблюдателя в другой системе отсчета K', движущейся относительно исходной системы K налево (рис. 11*б*) с малой (нерелятивистской) скоростью $-\mathbf{u}$. Так как $u \ll c$, то для преобразования скорости при переходе от K к K' можно

использовать классический закон сложения скоростей. Закон сохранения импульса требует, чтобы полный импульс тел до столкновения был равен импульсу образовавшегося тела. До столкновения полный импульс системы равен $2m\mathbf{u}$, где $m = m_0/\sqrt{1 - v^2/c^2}$ — релятивистская масса сталкивающихся тел; после столкновения он равен $M_0\mathbf{u}$, ибо вследствие $u \ll c$ массу образовавшегося тела и в K' можно считать равной массе покоя. Таким образом, из закона сохранения импульса следует, что масса покоя образовавшегося в результате неупругого соударения тела равна сумме релятивистских масс сталкивающихся частиц, т. е. она больше, чем сумма масс покоя исходных частиц:

$$M_0 = 2m = \frac{2m_0}{\sqrt{1-v^2/c^2}} > 2m_0.$$

Рассмотренный пример неупругого соударения двух тел, при котором происходит превращение кинетической энергии во внутреннюю энергию, показывает, что увеличение внутренней энергии тела также сопровождается пропорциональным увеличением массы. Этот вывод должен быть распространен на все виды энергии: нагретое тело имеет большую массу, чем холодное, сжатая пружина имеет большую массу, чем несжатая, и т. п.

Эквивалентность энергии и массы. Закон пропорциональности массы и энергии является одним из самых замечательных выводов теории относительности. Взаимосвязь массы и энергии заслуживает подробного обсуждения.

В классической механике масса тела есть физическая величина, являющаяся количественной характеристикой его инертных свойств, т. е. мера инертности. Это — инертная масса. С другой стороны, масса характеризует способность тела создавать поле тяготения и испытывать силу в поле тяготения. Это — тяготеющая, или гравитационная, масса. Инертность и способность к гравитационным взаимодействиям представляют собой совершенно различные проявления свойств материи. Однако то, что меры этих различных проявлений обозначаются одним и тем же словом, не случайно, а обусловлено тем, что оба свойства всегда существуют совместно и всегда друг другу пропорциональны, так что меры этих свойств можно выражать одним и тем же числом при надлежащем выборе единиц измерения.

Равенство инертной и гравитационной масс есть экспериментальный факт, подтвержденный с огромной степенью точности в опытах Этвеша, Дикке и др. Как же следует отвечать на вопрос: есть ли масса инертная и масса гравитационная одно и то же или нет? По своим проявлениям они различны, но их числовые характеристики пропорциональны друг другу. Такое положение вещей характеризуют словом «эквивалентность».

Аналогичный вопрос возникает в связи с понятиями массы покоя и энергии покоя в теории относительности. Проявления свойств материи, соответствующих массе и энергии, бесспорно различны. Но теория относительности утверждает, что эти свойства неразрывно связаны, пропорциональны друг другу. Поэтому в этом смысле можно говорить об *эквивалентности массы покоя и энергии покоя*. Выражающее эту эквивалентность соотношение (5) называется *формулой Эйнштейна*. Она означает, что всякое изменение энергии системы сопровождается эквивалентным изменением ее массы. Это относится к изменениям различных видов внутренней энергии, при которых масса покоя меняется.

О законе сохранения массы. Опыт показывает нам, что в громадном большинстве физических процессов, в которых изменяется внутренняя энергия, масса покоя остается неизменной. Как это согласовать с законом пропорциональности массы и энергии? Дело в том, что обычно подавляющая часть внутренней энергии (и соответствующей ей массы покоя) в превращениях не участвует и в результате оказывается, что определяемая из взвешивания масса практически сохраняется, несмотря на то, что тело выделяет или поглощает энергию. Это объясняется просто недостаточной точностью взвешивания. Для иллюстрации рассмотрим несколько численных примеров.

1. Энергия, высвобождающаяся при сгорании нефти, при взрыве динамита и при других химических превращениях, представляется нам в масштабах повседневного опыта громадной. Однако если перевести ее величину на язык эквивалентной массы, то окажется, что эта масса не составляет и 10^{-10} полной величины массы покоя. Например, при соединении 1 г водорода с 8 г кислорода выделяется около 10^{12} эрг энергии. Масса покоя образовавшейся воды на $\Delta m = \Delta E/c^2 \approx 10^{-9}$ г меньше массы исходных веществ. Такое изменение массы слишком мало для того, чтобы его можно было обнаружить с помощью современных приборов.

2. При неупругом столкновении двух частиц по 1 г, разогнанных навстречу друг другу до скорости 1 км/с, добавочная масса покоя слипшейся пары составляет

$$\Delta m \approx \left(2\cdot\frac{mv^2}{2}\right) : c^2 \approx 10^{-11}\ \text{г}.$$

(При такой скорости можно пользоваться нерелятивистским выражением для кинетической энергии.) Эта величина намного меньше погрешности, с которой может быть измерена масса 1 г.

Масса покоя и квантовые закономерности. Естественно задать вопрос: почему при обычных условиях подавляющая часть энергии находится в совершенно пассивном состоянии и в превращениях не участвует? На этот вопрос теория относительности не может дать ответа. Ответ следует искать в области квантовых закономерностей,

одной из характерных особенностей которых является существование устойчивых состояний с дискретными уровнями энергии.

Для элементарных частиц энергия, соответствующая массе покоя, либо превращается в активную форму (излучение) целиком, либо вовсе не превращается. Примером может служить превращение пары электрон—позитрон в гамма-излучение.

У атомов подавляющая часть массы находится в форме массы покоя элементарных частиц, которая в химических реакциях не изменяется. Даже в ядерных реакциях энергия, соответствующая массе покоя тяжелых частиц (нуклонов), входящих в состав ядер, остается пассивной. Но здесь активная часть энергии, т. е. энергия взаимодействия нуклонов, составляет уже заметную долю энергии покоя.

Таким образом, экспериментальное подтверждение релятивистского закона пропорциональности энергии покоя и массы покоя следует искать в мире физики элементарных частиц и ядерной физики. Например, в ядерных реакциях, идущих с выделением энергии, масса покоя конечных продуктов меньше массы покоя ядер, вступающих в реакцию. Соответствующая этому изменению массы энергия Δmc^2 с хорошей точностью совпадает с измеренной на опыте кинетической энергией образующихся частиц.

- Как импульс и энергия частицы зависят от ее скорости в релятивистской механике?
- Какая физическая величина называется массой частицы? Что такое масса покоя? Что такое релятивистская масса?
- Покажите, что релятивистское выражение (6) для кинетической энергии переходит в обычное классическое при $v \ll c$.
- Что такое энергия покоя? В чем заключается принципиальное отличие релятивистского выражения для энергии тела от соответствующего классического?
- В каких физических явлениях обнаруживает себя энергия покоя?
- Как понимать утверждение об эквивалентности массы и энергии? Приведите примеры проявления этой эквивалентности.
- Сохраняется ли масса вещества при химических превращениях?

△ **Вывод выражения для импульса.** Дадим обоснование формул (1), приведенных выше без доказательства, анализируя простой мысленный опыт. Для выяснения зависимости импульса частицы от скорости рассмотрим картину абсолютно упругого «скользящего» столкновения двух одинаковых частиц. В системе K_0 центра масс это столкновение имеет вид, показанный на рис. 12*а*: до столкновения частицы *1* и *2* движутся навстречу друг другу с одинаковыми по модулю скоростями, после столкновения частицы разлетаются в противоположные стороны с такими же по модулю скоростями, как и до столкновения. Другими словами, при столкновении

происходит только поворот векторов скоростей каждой из частиц на один и тот же угол θ.

Как будет выглядеть это же столкновение в других системах отсчета? Направим ось x вдоль биссектрисы угла θ и введем систему отсчета K, движущуюся вдоль оси x относительно системы центра масс со скоростью, равной x-составляющей скорости частицы *1*. В этой системе отсчета картина столкновения выглядит так, как показано на рис. 12*б*: частица *1* движется параллельно оси y, изменив при столкновении направление скорости и импульса на противоположное.

Сохранение y-составляющей полного импульса системы частиц при столкновении выражается соотношением

$$p_{1y} + p_{2y} = \tilde{p}_{1y} + \tilde{p}_{2y},$$

где $\tilde{p}_1$ и $\tilde{p}_2$ — импульсы частиц после столкновения. Так как $\tilde{p}_{1y} = -p_{1y}$ и $\tilde{p}_{2y} = -p_{2y}$ (рис. 12*б*), требование сохранения импульса означает равенство y-составляющих импульса частиц *1* и *2* в системе отсчета K:

$$p_{1y} = \tilde{p}_{2y}.$$

Теперь, наряду с K, введем в рассмотрение систему отсчета K', которая движется относительно системы K_0 центра масс со ско-

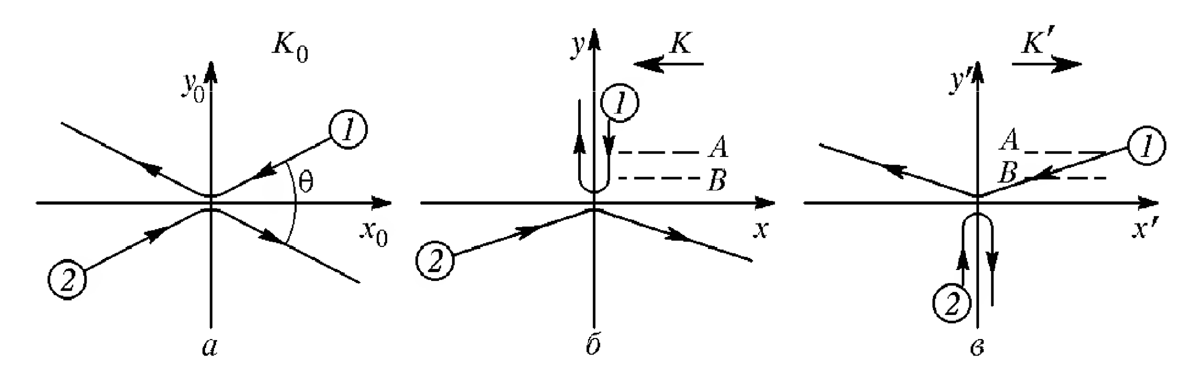


Рис. 12. К выводу зависимости массы тела от скорости

ростью, равной x-составляющей скорости частицы *2*. В этой системе K' частица *2* до и после столкновения движется параллельно оси y' (рис. 12*в*). Применяя закон сохранения импульса, убеждаемся, что в этой системе отсчета, как и в системе K, имеет место равенство y-составляющих импульса частиц *1* и *2*:

$$p'_{1y} = \tilde{p}'_{2y}.$$

Но из симметрии картин столкновения на рис. 12*б*, *в* легко сделать вывод о том, что модуль импульса частицы *1* в системе K равен модулю импульса частицы *2* в системе отсчета K', поэтому

$$p_{1y} = \tilde{p}'_{2y}.$$

Сопоставляя два последних равенства, находим $p_{1y} = p'_{1y}$, т. е. y-составляющая импульса частицы 1 одинакова в системах отсчета K и K'. Точно так же находим $p_{2y} = p'_{2y}$. Другими словами, y-составляющая импульса любой частицы, перпендикулярная направлению относительной скорости систем отсчета, одинакова в этих системах. В этом и заключается главный вывод из рассмотренного мысленного эксперимента.

Но y-составляющая скорости частицы имеет различное значение в системах отсчета K и K'. Согласно формулам преобразования скорости

$$u'_{1y} = u_{1y}\sqrt{1 - v^2/c^2},$$

где v есть скорость системы K' относительно K. Таким образом, в K' y-составляющая скорости частицы 1 меньше, чем в K.

Это уменьшение y-составляющей скорости частицы 1 при переходе от K к K' непосредственно связано с релятивистским преобразованием времени: одинаковое в K и K' расстояние между штриховыми линиями A и B (рис. 12*б*, *в*) частица 1 в системе K' проходит за большее время, чем в K. Если в K это время равно τ_0 (собственное время, так как оба события пересечение штрихов A и B — происходят в K при одном и том же значении координаты x), то в системе K' это время больше и равно $\tau = \tau_0/\sqrt{1 - v^2/c^2}$.

Вспоминая теперь, что y-составляющая импульса частицы 1 одинакова в системах K и K', мы видим, что в системе K', где y-составляющая скорости частицы меньше, этой частице нужно приписать как бы бóльшую массу, если под массой понимать, как и в нерелятивистской физике, коэффициент пропорциональности между скоростью и импульсом. Как уже отмечалось, этот коэффициент называют иногда релятивистской массой. Релятивистская масса частицы зависит от системы отсчета, т. е. является величиной относительной. В той системе отсчета, где скорость частицы много меньше скорости света, для связи между скоростью и импульсом частицы справедливо обычное классическое выражение $p = m_0 v$, где m_0 есть масса частицы в том смысле, как она понимается в нерелятивистской физике (масса покоя).

Будем считать, что в рассматриваемом нами «скользящем» столкновении скорость u_{1y} частицы 1 в системе K много меньше скорости света, т. е. ее масса в системе K есть масса покоя m_0, и $p_{1y} = m_0 u_{1y}$. Написав аналогичное выражение для y-составляющей импульса в системе K'

$$p'_{1y} = m u'_{1y},$$

где коэффициент пропорциональности, т. е. релятивистская масса частицы, обозначен буквой m, видим, что равенство $p_{1y} = p'_{1y}$ бу-

дет обеспечено, если коэффициенту m в системе отсчета K' приписать значение, даваемое формулой (3):

$$m = \frac{m_0}{\sqrt{1 - v^2/c^2}},$$

т. е. уменьшение поперечной составляющей скорости частицы при переходе от системы K к K' должно быть скомпенсировано возрастанием коэффициента пропорциональности m между скоростью и импульсом. Из приведенного вывода ясно, что это возрастание релятивистской массы, вызванное движением системы отсчета, действительно связано с релятивистским кинематическим эффектом замедления времени.

Возвращаясь к рис. 12, вспомним, что был рассмотрен случай скользящего столкновения, когда составляющая скорости частицы вдоль оси y была много меньше составляющей ее скорости вдоль оси x. В этом предельном случае входящая в полученную формулу относительная скорость v систем K и K' практически совпадает со скоростью частицы *1* в системе K'. Поэтому найденное значение коэффициента пропорциональности между y-составляющими векторов скорости и импульса справедливо и для самих векторов. Таким образом, соотношение (3) доказано.

Вывод выражения для энергии. Выясним теперь, к каким изменениям в выражении для энергии частицы приводит формула для релятивистского импульса.

В релятивистской механике сила $\mathbf{F}$ вводится таким образом, чтобы соотношение между приращением импульса частицы $\Delta\mathbf{p}$ и импульсом силы $\mathbf{F}\Delta t$ было таким же, как и в классической физике:

$$\Delta\mathbf{p} = \mathbf{F}\Delta t.$$

Будем считать, что энергия $E_к$ частицы в релятивистской механике, как и в классической, представляет собой величину, изменение которой на перемещении $\Delta\mathbf{r}$ равно работе действующей силы $\mathbf{F}$:

$$\Delta E_к = \mathbf{F}\Delta\mathbf{r} = \mathbf{F}\mathbf{v}\Delta t = \mathbf{v}\Delta\mathbf{p} = \mathbf{v}\Delta(m\mathbf{v}). \qquad (7)$$

Здесь перемещение частицы $\Delta\mathbf{r}$ за время Δt выражено через ее скорость $\mathbf{v}$. Из формулы (7) и будем исходить при выводе выражения для релятивистской энергии.

Перепишем формулу (3) следующим образом:

$$m^2(1 - v^2/c^2) = m_0^2.$$

Умножив обе части на c^2 и раскрыв скобки, получим

$$m^2c^2 - (m\mathbf{v})^2 = m^2c^2{}_0. \qquad (8)$$

При движении частицы под действием силы $\mathbf{F}$ ее скорость и импульс меняются. Для нахождения приращения левой части

(8) воспользуемся тем, что приращение квадрата любой переменной величины f за малый промежуток времени приближенно равно

$$\Delta(f^2) = (f + \Delta f)^2 - f^2 \approx 2f\Delta f.$$

Применяя эту формулу к равенству (8) и учитывая, что правая часть остается при этом неизменной, получаем

$$2mc\Delta(mc) - 2m\mathbf{v}\Delta(m\mathbf{v}) = 0,$$

откуда, после сокращения на $2m$, имеем

$$\Delta(mc^2) = \mathbf{v}\Delta(m\mathbf{v}). \tag{9}$$

Правые части в выражениях (7) и (9) совпадают. Поэтому левая часть (9) представляет собой приращение кинетической энергии частицы:

$$\Delta E_{\text{к}} = \Delta(mc^2). \tag{10}$$

Таким образом, приращение кинетической энергии частицы равно приращению ее релятивистской массы, умноженному на квадрат скорости света. Так как кинетическая энергия покоящейся частицы равна нулю, то из дифференциального соотношения (10), определяющего кинетическую энергию с точностью до константы, немедленно следует формула (6). ▲

- Как с помощью мысленного эксперимента можно показать, что составляющая импульса частицы, перпендикулярная направлению относительной скорости двух систем отсчета, одинакова в обеих этих системах? Какую роль при этом играют соображения симметрии?
- Поясните связь зависимости релятивистской массы частицы от ее скорости с релятивистским кинематическим эффектом замедления времени.
- Каким образом можно прийти к релятивистской формуле для кинетической энергии, основываясь на пропорциональности между приращениями кинетической энергии и релятивистской массы?

§ 5. Примеры релятивистского движения частиц

Формулы (1) предыдущего параграфа вместе с уравнениями (2), выражающими закон их изменения, служат основой для рассмотрения любых задач релятивистской механики. Вместо этих формул можно использовать две другие формулы, которые получаются из них исключением скорости $\mathbf{v}$:

$$E^2 - p^2c^2 = m_0^2c^4 \tag{1}$$

и почленным делением формул (1) § 4:

$$\mathbf{v} = \frac{c^2}{E}\,\mathbf{p}. \tag{2}$$

Связь энергии и импульса. Соотношение (1) устанавливает связь между энергией и импульсом частицы в релятивистской механике. Энергия и импульс частицы зависят от системы отсчета. Но правая часть в (1) представляет собой релятивистский инвариант. Это значит, что и стоящая в левой части комбинация $E^2 - p^2c^2$ изменяющихся при переходе от одной системы отсчета к другой энергии и импульса остается при таком переходе неизменной.

Для ультрарелятивистских частиц, энергия E которых много больше энергии покоя m_0c^2, соотношение (1) можно приближенно переписать в виде

$$E \approx pc \quad (\text{при } E \gg m_0c^2). \tag{3}$$

Формула (2) выражает скорость релятивистской частицы через ее энергию и импульс. В отличие от формул (1) § 4 эти формулы не теряют смысла даже при $v = c$ и поэтому пригодны для всех без исключения релятивистских объектов. Часто они оказываются более удобными и в практических приложениях.

Движение под действием постоянной силы. В качестве первого примера рассмотрим движение первоначально покоившейся частицы с зарядом q и массой покоя m_0 в однородном электрическом поле напряженности $\mathbf{E}$. Действующая на частицу сила $\mathbf{F}$ постоянна и равна $q\mathbf{E}$. Поэтому из закона изменения импульса $\Delta\mathbf{p} = \mathbf{F}\Delta t$ немедленно следует, что

$$\mathbf{p} = \mathbf{F}t. \tag{4}$$

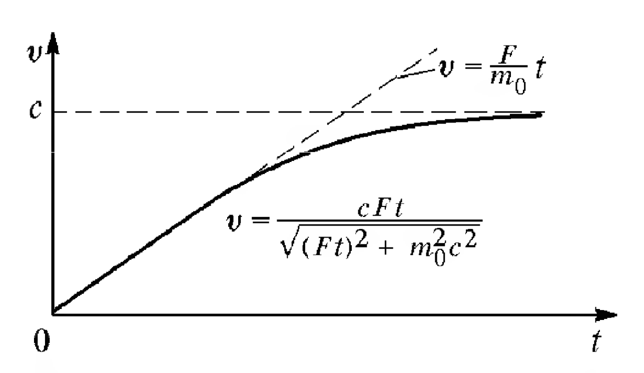


Рис. 13. Скорость частицы при движении в однородном электрическом поле

Подставляя это выражение для импульса частицы в формулу (1), получим

$$E^2 = (Ft)^2c^2 + m_0^2c^4. \tag{5}$$

Теперь с помощью (2) находим скорость частицы v спустя промежуток времени t после начала движения:

$$\mathbf{v} = \frac{\mathbf{F}t}{m_0\sqrt{1 + (Ft/m_0c)^2}}. \tag{6}$$

Если $Ft/m_0c \ll 1$, т. е. электрическое поле слабое или мало время движения, то в подкоренном выражении в (6) можно пренебречь вторым слагаемым, и для скорости получается обычное нерелятивистское выражение:

$$\mathbf{v} = \frac{\mathbf{F}}{m_0}t. \tag{7}$$

Если $Ft/m_0c \gg 1$, то под корнем в (6) можно пренебречь единицей по сравнению со вторым слагаемым. Видно, что при этом скорость v стремится к c. На рис. 13 показана зависимость скорости v от времени.

Движение в магнитном поле. Перейдем к рассмотрению движения заряженной частицы в однородном магнитном поле. Поскольку действующая на частицу со стороны магнитного поля сила Лоренца перпендикулярна скорости частицы, то скорость не меняется по модулю и, следовательно, не меняется и релятивистская масса частицы m. Поэтому закон изменения импульса частицы запишется в виде

$$m\frac{d\mathbf{v}}{dt} = q\mathbf{v}\times\mathbf{B}. \tag{8}$$

Если скорость $\mathbf{v}$ перпендикулярна вектору индукции магнитного поля $\mathbf{B}$, то частица движется по окружности и ее ускорение равно v^2/R, где R — радиус окружности. В этом случае уравнение (8) дает

$$m\frac{v^2}{R} = qvB. \tag{9}$$

Для угловой скорости обращения ω_c, связанной с линейной скоростью v обычным соотношением $v = \omega_c R$, с помощью (9) находим

$$\omega_c = \frac{qB}{m}. \tag{10}$$

Выражение (10) имеет такой же вид, как и нерелятивистская формула для угловой скорости обращения в магнитном поле, только в знаменателе стоит релятивистская масса частицы m, связанная с ее массой покоя m_0 соотношением

$$m = \frac{m_0}{\sqrt{1-v^2/c^2}}.$$

Ускоритель на встречных пучках. В качестве третьего примера рассмотрим ускоритель заряженных частиц на встречных пучках. Выясним, в чем преимущество таких ускорителей по сравнению с обычными ускорителями с неподвижной мишенью, и установим соответствие между кинетической энергией частицы $E_к$ в обычном ускорителе и эквивалентной энергией $E'_к$ в ускорителе со встречными пучками.

Одной из важнейших характеристик ускорителя является та доля кинетической энергии разогнанных элементарных частиц, которая может быть использована для реакции образования новых частиц. В обычных ускорителях, когда частица-мишень неподвижна, требование сохранения импульса исключает возможность превращения всей кинетической энергии частицы-снаряда в энергетический эквивалент массы покоя новых частиц, образующихся при столкновении. В самом деле, до столкновения суммарный импульс снаряда и мишени отличен от нуля. Такой же суммарный импульс должен

быть и после столкновения. Поэтому образовавшиеся в результате столкновения частицы не могут находиться в покое и, следовательно, часть начальной кинетической энергии снаряда переходит в кинетическую энергию частиц после столкновения.

Однако если сталкивающиеся частицы с равными массами летят навстречу друг другу с одинаковыми скоростями, то в результате неупругого удара вся кинетическая энергия налетающих частиц может быть использована для рождения новых частиц: поскольку начальный импульс системы равен нулю, то ничто не запрещает покоиться образовавшимся в результате столкновения частицам.

Оценим вначале «выигрыш» в энергии для простого случая столкновения одинаковых нерелятивистских частиц. Используя закон сохранения импульса, легко убедиться, что в этом случае при неподвижной мишени для реакции образования новых частиц может быть использована только половина кинетической энергии налетающей частицы $E_к$: $\Delta E = E_к/2$. Если же столкнутся движущиеся навстречу друг другу частицы с кинетическими энергиями $E_к$, то для реакции может быть использована вся их кинетическая энергия $\Delta E = 2E_к$. Таким образом, используя ускоритель, способный разогнать частицы до кинетической энергии $E_к$, мы можем с помощью накопительных колец повысить эффективность использования кинетической энергии в четыре раза.

Идея устройства накопительных колец показана на рис. 14. Пучок частиц из ускорителя с помощью быстродействующего магнита-переключателя A разделяется на два пучка, каждый из которых с помощью системы отклоняющих магнитов A и B направляется в свое кольцо, где обращается по орбите благодаря удерживающему магнитному полю, перпендикулярному плоскости рисунка. На общем участке CD происходят столкновения движущихся навстречу друг другу частиц.

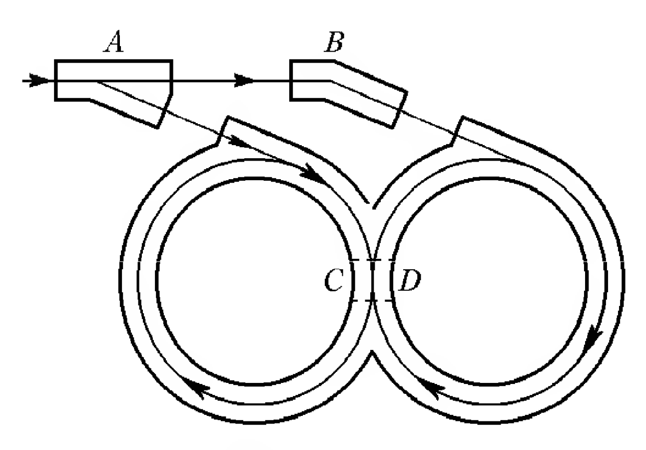


Рис. 14. Накопительные кольца

Итак, в нерелятивистском случае неупругого столкновения частиц одинаковой массы, одна из которых покоится, т. е. при использовании неподвижной мишени, только половина первоначальной энергии может перейти в энергию покоя рождающихся частиц. А как обстоит дело в случае релятивистских частиц, с которыми имеет дело физика высоких энергий? Оказывается, что для неподвижной мишени дело обстоит еще хуже. Чтобы убедиться в этом, придется тщательно рассмотреть законы сохранения энергии и импульса при столкновении релятивистских частиц.

Рассмотрим неупругое столкновение релятивистской частицы с массой покоя m_0 с такой же покоящейся частицей.

Будем искать энергию ΔE, которая может быть использована для образования новых частиц в этом случае. Обозначим через M_0 полную массу покоя системы после столкновения. Тогда ΔE есть не что иное, как увеличение энергии покоя частиц, которое произошло в рассматриваемом столкновении:

$$\Delta E = M_0c^2 - 2m_0c^2. \tag{11}$$

Найдем теперь M_0 — массу покоя частиц системы после столкновения. Применим к столкновению законы сохранения энергии и импульса. Из формулы (1) выразим квадрат импульса любой частицы p^2 через ее полную энергию E:

$$p^2 = \frac{E^2}{c^2} - m_0^2c^2. \tag{12}$$

Полная энергия релятивистской частицы E есть сумма энергии покоя частицы и ее кинетической энергии:

$$E = 2m_0c^2 + E_к. \tag{13}$$

Энсргия, которой характсризуют ускоритсли, — это кинстичсская энергия разогнанных частиц $E_к$. Учитывая, что до столкновения одна из частиц покоилась ($p_2 = 0$), запишем квадрат импульса всей системы до удара P^2, равный квадрату импульса налетающей частицы p_1^2 (12), в виде

$$P^2 = \left(\frac{m_0c^2 + E_к}{c}\right)^2 - m_0^2c^2. \tag{14}$$

Согласно закону сохранения энергии полная энергия системы после столкновения E' такая же, как и до столкновения, т. е. равна сумме энергий покоя обеих частиц и кинетической энергии налетающей частицы:

$$E' = m_0c^2 + E_к. \tag{15}$$

Запишем теперь квадрат импульса системы после столкновения с помощью (12) и (15):

$$P^2 = \frac{E'^2}{c^2} - M_0^2c^2 = \left(\frac{2m_0c^2 + E_к}{c}\right)^2 - M_0^2c^2. \tag{16}$$

Полный импульс системы до удара (14) и после удара (16) обозначены одной и той же буквой P, так как полный импульс системы сохраняется. Приравнивая правые части равенств (14) и (16), после простых преобразований находим M_0:

$$M_0 = 2m_0\left(1 + \frac{E_к}{2m_0c^2}\right)^{1/2}. \tag{17}$$

Теперь для ΔE в (11) получим

$$\Delta E = 2m_0c^2\left(\sqrt{1+\frac{E_\text{к}}{2m_0c^2}}-1\right). \tag{18}$$

Легко видеть, что для нерелятивистской частицы, кинетическая энергия которой много меньше энергии покоя $E_\text{к} \ll m_0c^2$, выражение (18) дает результат, полученный ранее элементарным путем: $\Delta E = E_\text{к}/2$. Для этого достаточно воспользоваться приближенной формулой

$$\sqrt{1+x} \approx 1 + x/2 \quad \text{при} \quad x \ll 1.$$

В противоположном ультрарелятивистском случае, когда кинетическая энергия частицы много больше энергии покоя: $E_\text{к} \gg m_0c^2$, в формуле (18) можно пренебречь единицами по сравнению с $E_\text{к}/(2m_0c^2)$. Тогда выражение (18) принимает вид

$$\Delta E \approx \sqrt{E_\text{к} \cdot 2m_0c^2}. \tag{19}$$

Если, например, мы хотим иметь $\Delta E = 20$ ГэВ при столкновении протонов (энергия покоя протона $2m_0c^2 \approx 1$ ГэВ), то с помощью формулы (19) убеждаемся, что необходим ускоритель, разгоняющий протоны до энсргии $E_\text{к} \approx 200$ ГэВ. Таким образом, в рассматривасмом примере может быть использована только десятая часть кинетической энергии протона (а не половина, как было бы в нерелятивистском случае).

Итак, из-за релятивистских эффектов доля кинетической энергии разогнанных частиц, которая может быть использована для реакции, у ускорителей с неподвижной мишенью падает с ростом энергии. В ускорителе же на встречных пучках и в релятивистском случае вся кинетическая энергия сталкивающихся частиц может перейти в энергию покоя рождающихся частиц.

Интересно получить соотношение, связывающее кинетические энергии частиц в ускорителе обычного типа $E_\text{к}$ и в ускорителе на встречных пучках $E'_\text{к}$, при которых получается одна и та же энергия ΔE, способная превратиться в энергию покоя рождающихся частиц. В ускорителе на встречных пучках $\Delta E = 2E'_\text{к}$. В ускорителе с неподвижной мишенью ΔE определяется формулой (18). Подставляя в нее $\Delta E = 2E'_\text{к}$, находим

$$E_\text{к} = 2m_0c^2\left[\left(\frac{E'_\text{к}}{m_0c^2}+1\right)^2-1\right]. \tag{20}$$

В ультрарелятивистском случае, когда $E'_\text{к} \gg m_0c^2$, эта формула принимает вид

$$E_\text{к} \approx \frac{2E'^2_\text{к}}{m_0c^2}. \tag{21}$$

Из приведенных формул видно, что выигрыш при использовании ускорителей на встречных пучках особенно велик для легких частиц, например электронов, для которых $m_0c^2 \approx 0{,}5$ МэВ. Так, для установки со встречными пучками, ускоряющей электроны до энергии $E'_\text{к} = 130$ МэВ, энергия $E_\text{к}$ эквивалентного ускорителя с неподвижной мишенью составляет, согласно формуле (21), примерно 70 ГэВ, т. е. в 520 раз больше!

- При каких условиях энергия частицы пропорциональна импульсу?
- Покажите, что при ускорении частицы под действием постоянной силы ее скорость стремится к конечному пределу.
- Почему в ускорителях, разгоняющих частицы до высоких энергий, период обращения в магнитном поле не остается постоянным?
- В чем заключается преимущество использования встречных пучков в физике высоких энергий?

II. ЗАКОНЫ МИКРОМИРА. ЧАСТИЦЫ И ВОЛНЫ

§ 6. Световые кванты

Одна из величайших революций в физике, связанная с открытием квантовых закономерностей, пришлась на начало XX столетия. Оказалось, что ряд экспериментальных фактов, относящихся к атомным явлениям, в принципе невозможно объяснить в рамках классической механики и электродинамики, получивших к этому времени вполне законченный вид.

Фотоэлектрический эффект. Наиболее ярко ограниченность представлений классической физики обнаружила себя в явлении фотоэффекта, открытого Г. Герцем и впервые подробно исследованном А. Г. Столетовым.

Явление фотоэффекта заключается в вырывании электронов из вещества падающим на него светом. Основные черты этого явления сводятся к следующему. Пучок света, падающий на поверхность металла, освобождает из металла электроны при условии, что частота света выше определенного критического значения, зависящего от рода металла. Количество вырываемых в единицу времени электронов при неизменном спектральном составе излучения пропорционально падающему на поверхность металла световому потоку.

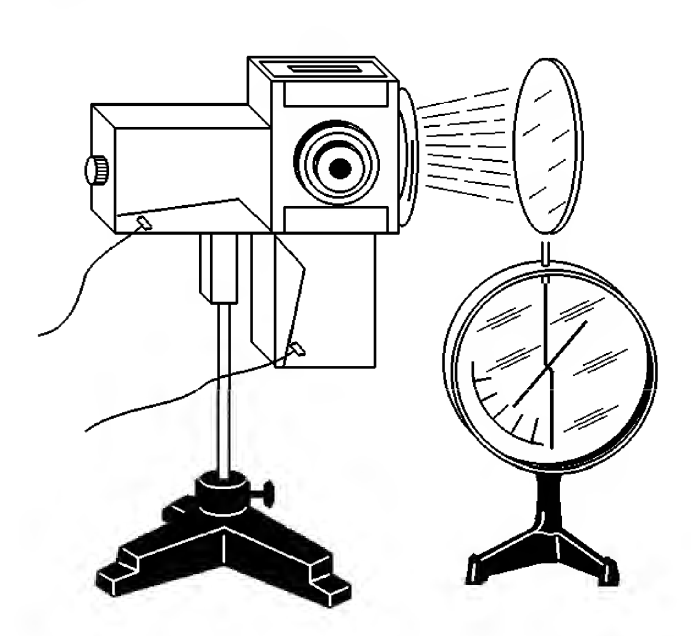


Рис. 15. Простейший опыт по наблюдению фотоэффекта

К моменту открытия фотоэффекта в 1887 г. еще ничего не было известно об электронах, открытых Дж. Томсоном только десять лет спустя в 1897 г. Однако экспериментально было установлено, что металл при освещении теряет именно отрицательный заряд.

Экспериментальные закономерности фотоэффекта. Простейший способ наблюдать фотоэффект — это освещать ультрафиолетовым излучением заряженную цинковую пластину, соединенную с электрометром (рис. 15). Если заряд пластины положительный, то облучение ее никак не влияет на быстроту разрядки электрометра. Но если заряд пластины отрицательный, то излучение электрической дуги очень быстро разряжает электрометр. Сразу после открытия электрона стало ясно, что фотоэффект связан именно с освобождением имеющихся в металле электронов. Если электроны, испускае-

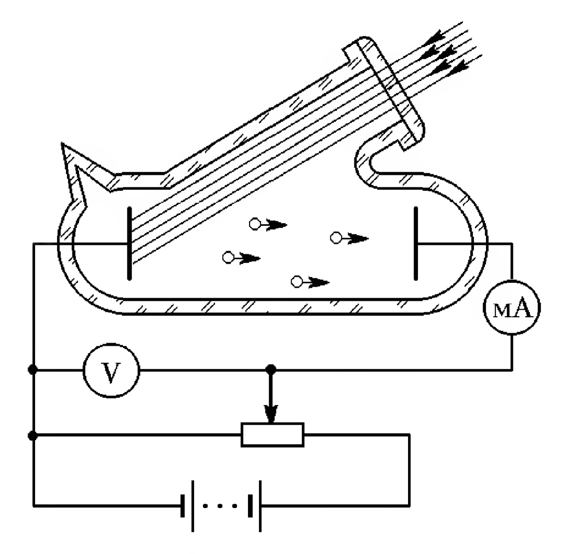


Рис. 16. Схема экспериментальной установки для изучения фотоэффекта

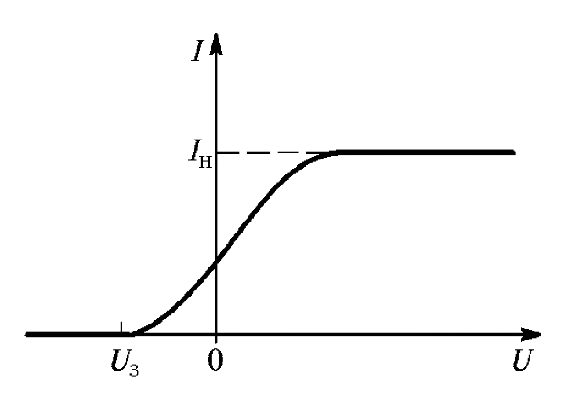


Рис. 17. Зависимость силы фототока от приложенного напряжения при неизменном световом потоке

мые одним электродом (их обычно называют фотоэлектронами), собрать на другом электроде в вакуумной колбе и замкнуть между электродами цепь (рис. 16), то сила тока в ней будет пропорциональна световому потоку. Весь процесс протекает практически мгновенно, т. е. *безынерционно.*

Если, не меняя светового потока, увеличивать разность потенциалов между электродами, то сила тока (фототока) нарастает. При некотором напряжении она достигает определенного максимального значения, после чего перестает увеличиваться (рис. 17). Это максимальное значение фототока называется *током насыщения.* Ясно, что насыщение фототока наступает тогда, когда все испускаемые фотоэлектроны попадают на второй электрод.

Освобожденные из металла светом электроны обладают некоторой кинетической энергией. Эта энергия возрастает с увеличением частоты падающего света.

Кинетическая энергия фотоэлектронов. Остановимся на том, как можно измерить кинетическую энергию фотоэлектронов. Из графика на рис. 17 видно, что сила фототока отлична от нуля и при отсутствии приложенного напряжения. Если изменить полярность батареи, то ток уменьшится и при некотором напряжении $U_з$ сила то-

ка обратится в нуль. Это так называемое *задерживающее напряжение*, при котором все вырванные светом электроны возвращаются на катод, из которого они были вырваны. Значение $U_з$ зависит от максимальной кинетической энергии, которую имеют вырванные светом электроны. Измеряя задерживающее напряжение и применяя закон сохранения энергии, можно найти это максимальное значение кинетической энергии:

$$\frac{mv^2_{\text{макс}}}{2} = eU_з.$$

При изменении интенсивности света сила тока насыщения $I_н$ также изменяется, но, как показали опыты, задерживающее напряжение $U_з$ остается неизменным. Это означает, что максимальная кинетическая энергия фотоэлектронов не зависит от интенсивности света (при его неизменном спектральном составе). С точки зрения волновых представлений о свете этот факт необъясним. Ведь чем больше интенсивность света, тем большие силы действуют на электроны в освещенном металле и тем большая энергия должна, казалось бы, передаваться светом электронам.

Тщательные опыты показали, что кинетическая энергия электронов, испускаемых поверхностью металла, освещаемой видимым или ультрафиолетовым светом, не зависит от интенсивности излучения, а зависит лишь от рода металла и от частоты излучения ν. Теоретическое объяснение наблюдаемой закономерности было дано в 1905 г. А. Эйнштейном на основе гипотезы Планка о том, что свет излучается определенными порциями, причем энергия каждой такой порции определяется формулой $E = h\nu$.

Соотношение Планка. Эта формула была предложена Планком при разработке теории теплового излучения. Последовательное применение классической электродинамики к этой проблеме приводило к физически бессмысленному выводу о том, что спектральная плотность излучения нагретого тела должна неограниченно возрастать в области высоких частот. Этот парадоксальный вывод из классических представлений получил образное название «ультрафиолетовой катастрофы».

Такое противоречие теории и эксперимента не возникает, если предположить, что вещество испускает электромагнитную энергию не непрерывно, а отдельными порциями — *квантами*, энергия которых пропорциональна частоте ν излучения:

$$E = h\nu.$$

Коэффициент пропорциональности h получил название *постоянной Планка*. Числовое значение постоянной Планка равно $6{,}626 \cdot 10^{-34}$ Дж·с. Построенная Планком теория теплового излучения превосходно согласуется с экспериментом, однако положенное в осно-

ву этой теории предположение означает, что законы классической физики неприменимы к описанию явления теплового излучения.

Уравнение Эйнштейна. Описывающее фотоэффект *уравнение Эйнштейна* выражает закон сохранения энергии в элементарном акте взаимодействия светового кванта, имеющего энергию $h\nu$, с электроном металла:

$$h\nu = \frac{mv^2_{\text{макс}}}{2} + A. \qquad (1)$$

В этом выражении A — так называемая *работа выхода* электрона из металла, т. е. минимальная энергия, необходимая для удаления электрона из металла, $mv^2_{\text{макс}}/2$ — максимальная кинетическая энергия освобождаемого электрона.

Из уравнения (1) можно сделать вывод, что свет с частотой ν, меньшей некоторого значения $\nu_{\text{мин}}$, определяемого соотношением $h\nu_{\text{мин}} = A$, вообще не может вызывать фотоэффект в веществе с данным значением работы выхода, так как энергия $h\nu$ кванта оказывается меньше работы выхода. Эта частота $\nu_{\text{мин}}$ определяет так называемую *красную границу* фотоэффекта. У цинка красная граница фотоэффекта приходится на ближнюю ультрафиолетовую область спектра. У некоторых других металлов, например у натрия и калия, эта граница лежит в видимой области спектра.

Детальное изучение опытных данных показало, что h в выражении (1) действительно совпадает с универсальной постоянной (постоянной Планка) $h = 6{,}626 \cdot 10^{-34}$ Дж·с, значение которой впервые было найдено из опытов по спектральному составу теплового излучения.

Корпускулярные и волновые свойства света. Попытаемся понять, почему электрон обычно поглощает только один квант энергии независимо от интенсивности излучения[1]. Самое простое объяснение этого явления заключается в следующем: свет состоит из частиц, т. е. из излученных порций световой энергии $E = h\nu$, которые сохраняют свою индивидуальность в процессе распространения и в дальнейшем, при столкновении с электроном, передают ему всю свою энергию.

Это предположение подкрепляется различными опытами, например такими, когда на фотопластинку направляются лучи очень низкой интенсивности. На пластинке получаются хаотически разбросанные темные точки, распределенные со средней плотностью, пропорциональной интенсивности света. Для очень интенсивного пучка распределение точек становится настолько плотным, что они практически сливаются. При такой интенсивности пучка света дискретность

[1] При очень высокой интенсивности сфокусированного лазерного излучения возможен многофотонный фотоэффект, когда одному электрону передается энергия сразу нескольких фотонов. Для многофотонного фотоэффекта красная граница отсутствует.

актов взаимодействия света с фотопластинкой уже не проявляется, и пучок становится похожим на то, что в классической физике называется световой волной.

Хотя предположение, что свет состоит из локализованных частиц, позволяет просто объяснить фотоэлектрический эффект, оно не согласуется с огромной совокупностью экспериментов, приводящих к выводу, что свет является формой волнового движения. Вспомним чередование интенсивности дифракционных полос света, падающего на экран после прохождения через одну или несколько щелей. Часто имеет место такое явление, что при двух открытых щелях, близких друг к другу, интенсивность света в некоторых местах экрана будет очень мала, в то время как в тех же местах при пропускании света только через одну щель наблюдается высокая интенсивность. Эти результаты легко объясняются предположением, что свет представляет собой волны, которые, интерферируя, могут усиливаться или ослабляться. Но их совершенно невозможно объяснить, если предположить, что свет состоит из классических частиц-корпускул. Такие частицы должны были бы проходить или через одну, или через другую щель, и наличие второй щели едва ли могло бы влиять на характер движения частиц, проходящих через первую щель.

Невозможность классического объяснения фотоэффекта. Можно ли объяснить закономерности фотоэффекта исходя из волновых представлений о свете? Попробуем рассмотреть фотоэффект с классической точки зрения. Взаимодействуя с электроном, находящимся внутри атома, излучение передает ему свою энергию. Электрон будет поглощать энергию световой волны, пока он не освободится из атома. Объяснить уравнение (1) можно, предположив, что свойства атома таковы, что электрон будет сохранять полученную от света энергию и находиться в атоме до тех пор, пока не накопит ее до величины $h\nu$, после чего покидает атом. Если бы атом действительно обладал такими свойствами, то для света с очень маленькой интенсивностью фотоэффект не должен был бы наблюдаться в течение очень долгого времени, так как должно было бы пройти значительное время, чтобы накопилась необходимая порция энергии. Соответствующие опыты проводились с металлическими пылинками и очень слабым светом. Пылинки были настолько малы, что потребовалось бы много часов для накопления энергии $h\nu$. Однако немедленно после начала их освещения появлялось некоторое количество фотоэлектронов. Итак, эта попытка объяснения не удалась.

Такова же была судьба всех остальных подобных попыток объяснения закономерностей фотоэффекта на основе волновых представлений. Это означает, что волновая теория не способна объяснить внезапную локализацию конечных порций энергии на одной классической частице.

Двойственная природа света. Обобщение данных различных опытов привело к следующим не известным ранее важным выводам: во-первых, к представлению о двойственном характере электромагнитного излучения, проявляющего то волновые, то корпускулярные свойства, и, во-вторых, к утверждению о существовании дискретных значений некоторых из тех физических величин, которые, по представлениям классической физики, могли меняться непрерывно.

Итак, одна группа опытов указывает, что свет — это частицы, которые могут быть локализованы; другая же группа не менее убедительных опытов доказывает, что свет — это волны. Какое же из этих утверждений правильно? Вопрос стоит так: можно ли найти единое представление, объясняющее все наши сведения о природе света? Позже мы вернемся к обсуждению этого вопроса, а пока будем просто считать, что свет существует в форме квантов-фотонов, которые при некоторых условиях проявляют себя подобно частицам, а при других — подобно волнам.

Фотоны. Мы уже подробно рассматривали волновые свойства света. Теперь рассмотрим более детально свет с корпускулярной точки зрения. В теории Планка энергия фотона E связана с частотой света ν соотношением

$$E = h\nu. \tag{2}$$

Согласно теории относительности энергия всегда связана с массой соотношением

$$E = mc^2.$$

Поэтому релятивистская масса фотона

$$m = \frac{h\nu}{c^2}. \tag{3}$$

Поскольку фотоны не существуют в состоянии покоя, то их масса покоя $m_0 = 0$, а масса m, определяемая формулой (3), — это масса фотона, движущегося в вакууме со скоростью c. Импульс фотона равен произведению его релятивистской массы на скорость:

$$p = mc = \frac{h\nu}{c} = \frac{E}{c}. \tag{4}$$

Отметим, что формула (4), дающая связь между энергией фотона E и его импульсом p, является частным случаем общего соотношения, связывающего энергию, импульс и массу покоя любой частицы:

$$\left(\frac{E}{c}\right)^2 = m_0^2c^2 + p^2.$$

Эффект Комптона. Опыт показывает, что свет может взаимодействовать с веществом только путем дискретных процессов, при которых испускается или поглощается целый квант. С корпускулярной точки зрения взаимодействие между веществом и светом описывается как поглощение, испускание или рассеяние фотонов, сопровождающееся изменением их энергии и импульса. Экспериментально рассеяние фотонов на электронах было исследовано А. Комптоном в 1923 г. В его опытах через вещество с легкими атомами (графит, парафин) пропускался пучок рентгеновских лучей частоты ν. Измерения Комптона показали, что в рассеянном рентгеновском свете, наряду с излучением неизменной длины волны, появляется рентгеновское излучение с несколько большей длиной волны. Наблюдаемое изменение длины волны $\Delta\lambda = \lambda' - \lambda$ зависит от угла θ между направлением первичного пучка и направлением рассеянного света следующим образом:

$$\Delta\lambda = 2k \sin^2 \frac{\theta}{2}, \qquad (5)$$

причем постоянная k, найденная из эксперимента, равна 0,0024 нм.

Законы сохранения в эффекте Комптона. Для объяснения эффекта Комптона применим законы сохранения энергии и импульса к столкновению рентгеновского фотона с электроном, как если бы фотон представлял собой обычную частицу. В атомах легких элементов для удаления электрона нужна энергия порядка 10 эВ, что примерно в тысячу раз меньше энергии рентгеновского кванта $h\nu \approx 10$ кэВ. Поэтому электроны в этих опытах можно считать практически свободными. Энергия покоя электрона $m_0c^2 = 0{,}5$ МэВ, и отношение $h\nu/(m_0c^2) \ll 1$. Следовательно, покоившийся до столкновения с фотоном или двигавшийся в атоме с нерелятивистской скоростью электрон и после столкновения останется нерелятивистским. Закон сохранения энергии при столкновении в пренебрежении начальной энергией электрона записывается в виде

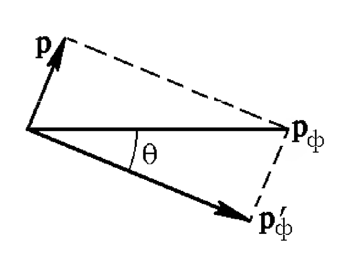


Рис. 18. К объяснению эффекта Комптона

$$h\nu = h\nu' + \frac{p^2}{2m_0}, \qquad (6)$$

где p — импульс электрона после столкновения с фотоном. Закон сохранения импульса

$$\mathbf{p}_ф = \mathbf{p}'_ф + \mathbf{p}, \qquad (7)$$

где $\mathbf{p}_ф$ и $\mathbf{p}'_ф$ — импульсы фотона до и после рассеяния, запишем с помощью теоремы косинусов (рис. 18) в виде

$$p^2 = p_ф^2 + p_ф'^2 - 2p_ф p'_ф \cos\theta. \qquad (8)$$

Подставляем в это равенство выражение для импульса фотона через его частоту $p_ф = h\nu/c$ и квадрат импульса электрона из закона сохранения энергии (6):

$$2m_0 h(\nu - \nu') = \frac{h^2}{c^2}\left(\nu^2 + \nu'^2 - 2\nu\nu' \cos\theta\right).$$

Переписав это уравнение в виде

$$\nu - \nu' = \frac{h\nu}{2m_0 c^2}\left(\nu + \frac{\nu'^2}{\nu} - 2\nu' \cos\theta\right) \tag{9}$$

и учитывая, что $h\nu/(m_0 c^2) \ll 1$, видим, что изменение частоты $\Delta\nu = \nu - \nu'$ мало по сравнению с самой частотой ν. Поэтому в правой части (9) можно ν' заменить на ν. Тогда для относительного сдвига частоты $\Delta\nu/\nu$ при рассеянии сразу получаем

$$\frac{\Delta\nu}{\nu} = -\frac{h\nu}{m_0 c^2}(1 - \cos\theta). \tag{10}$$

Знак «—» показывает что частота рентгеновского излучения при рассеянии уменьшается. Это естественно, поскольку фотон отдает часть своей энергии электрону. Теперь перейдем от частот к длинам волн. Это легко сделать, учитывая, что относительное изменение частоты при рассеянии мало: $\Delta\nu/\nu \ll 1$. Тогда

$$\lambda + \Delta\lambda = \frac{c}{\nu + \Delta\nu} \approx \frac{c}{\nu}\left(1 - \frac{\Delta\nu}{\nu}\right).$$

Так как $c/\nu = \lambda$, то из этого выражения имеем $\Delta\lambda = -\lambda\Delta\nu/\nu$, и соотношение (10) переписывается в виде

$$\Delta\lambda = \frac{h}{m_0 c}(1 - \cos\theta) = \frac{2h}{m_0 c}\sin^2\frac{\theta}{2}. \tag{11}$$

Эта формула совпадает с (5), ибо величина $h/(m_0 c)$, как легко убедиться, как раз равна 0,0024 нм. Эта имеющая размерность длины волны комбинация трех универсальных постоянных получила название *комптоновской длины волны* электрона.

Интересно отметить, что изменение длины волны рентгеновского излучения в явлении Комптона, как видно из (11), не зависит от длины волны падающего излучения. А как объяснить существование в спектре рассеянного излучения еще и несмещенной линии? Все дело в том, что внутренние электроны, особенно в тяжелых атомах, связаны настолько прочно, что энергия, необходимая для их удаления, уже сравнима с энергией рентгеновских квантов и, следовательно, внутренние электроны нельзя рассматривать как свободные. Поэтому при соударении фотон обменивается энергией и импульсом с атомом в целом, а так как масса атома очень велика, то по закону сохранения импульса фотон практически не передает ему своей энергии. Следовательно, энергия кванта $h\nu$ при таком рассеянии не изменяется. Гамма-кванты, рассеянные внутренними электронами,

образуют несмещенную компоненту, а внешними — смещенную. Из приведенных рассуждений ясно, почему эффект Комптона нельзя наблюдать в видимой области спектра. Дело в том, что энергия фотона видимого света составляет лишь несколько электрон-вольт (1 эВ $= 1{,}6\cdot 10^{-19}$ Дж $= 1{,}6\cdot 10^{-12}$ эрг). Для такого света даже внешние электроны в легких атомах нельзя считать свободными.

Дискретный характер взаимодействия света с электронами. Опыты Комптона ярко продемонстрировали, что энергия и импульс фотона действительно выражаются формулами (2) и (4), а также то, что законы сохранения энергии и импульса выполняются при индивидуальных процессах рассеяния. Законы сохранения могут быть проверены еще более полно, если исследовать электроны отдачи. Соответствующие измерения показали, что при этом электроны приобретают те же самые импульс и энергию, которые теряет фотон.

Как и в случае фотоэффекта, объяснить особенности явления Комптона исходя из волновой точки зрения на свет не удается. Взаимодействуя с электромагнитной волной, классический электрон мог бы получать любое количество энергии. В спектре рассеянного излучения при наблюдении под заданным углом θ можно было бы обнаружить различные значения для изменения длины волны при изменении интенсивности излучения или времени экспозиции. Однако в опытах было однозначно показано, что при заданном значении угла наблюдается только одно значение смещения длины волны независимо от интенсивности излучения и времени облучения. Эти факты указывают на то, что процесс передачи энергии и импульса является не непрерывным, как это предсказывает классическая теория, а дискретным.

Корпускулярное объяснение эффекта Доплера. Но не всегда вопрос о природе света стоит так категорично: или кванты, или классические волны. Существует ряд явлений, допускающих корректное объяснение с любой из этих точек зрения. Таков, например, рассмотренный ранее эффект Доплера. Это типично волновое, на первый взгляд, явление может быть объяснено с точки зрения представления о свете как о потоке фотонов. Покажем это. Пусть «закрепленный» неподвижный атом испускает фотон с энергией $h\nu$ при переходе атома из одного состояния с определенной энергией в другое. Разность энергий этих состояний $h\nu$ не зависит от того, покоится атом или движется.

При испускании фотона свободно движущимся атомом импульс атома изменяется, поскольку испущенный фотон обладает импульсом. Следовательно, кинетическая энергия атома также изменяется. Энергия фотона $h\nu'$, испущенного движущимся атомом, отличается от $h\nu$ вследствие изменения кинетической энергии атома.

На основании закона сохранения энергии

$$h\nu' - h\nu = \frac{\mathbf{p}^2}{2M} - \frac{\mathbf{p}_1^2}{2M}, \tag{12}$$

где $\mathbf{p}$ — импульс атома до испускания фотона, $\mathbf{p}_1$ — после испускания, M — масса атома.

Начальный и конечный импульсы атома можно связать с импульсом испускаемого фотона $\mathbf{p}_ф$ с помощью закона сохранения импульса (рис. 19):

$$\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_ф. \tag{13}$$

Рис. 19. К объяснению эффекта Доплера при испускании фотона движущимся атомом

Перенося $\mathbf{p}_ф$ в равенстве (13) в левую часть, возводя полученное равенство в квадрат и учитывая, что импульс фотона крайне мал по сравнению с импульсом излучающего атома, получаем

$$p^2 - 2pp_ф \cos\theta \approx p_1^2. \tag{14}$$

С помощью (14) соотношение (12) можно переписать в виде

$$h\nu' - h\nu = \frac{p}{M}\, p_ф \cos\theta. \tag{15}$$

Подставляя в (15) импульс испущенного фотона $p_ф = h\nu'/c$ и учитывая, что p/M есть скорость движения v излучающего атома, находим

$$\nu' - \nu = \nu' \frac{v}{c} \cos\theta,$$

откуда

$$\frac{\Delta\nu}{\nu} = \frac{v}{c}\cos\theta \tag{16}$$

с точностью до членов порядка v/c. Это есть обычное, нерелятивистское выражение для явления Доплера.

Изложенные здесь факты говорят о том, что основанные на классической физике представления о природе света требуют радикального пересмотра.

- Перечислите основные экспериментальные закономерности фотоэффекта. Какие из них вступают в противоречие с основными представлениями классической физики?
- Объясните физический смысл уравнения Эйнштейна для фотоэффекта. Какова связь этого уравнения с законом сохранения энергии?
- Как связана энергия кванта света с его частотой и с длиной волны?
- Какие из известных вам опытов свидетельствуют о волновых свойствах света?
- Какие опыты со светом не допускают объяснения на основе волновых представлений?

- В чем проявляется двойственный характер света?
- Как связаны между собой энергия и импульс фотона?
- Как проявляются корпускулярные свойства электромагнитного излучения в явлении Комптона?
- На чем основана корпускулярная трактовка явления Доплера?
- Какие явления свидетельствуют о том, что законы сохранения энергии и импульса выполняются в элементарных актах взаимодействия света с веществом?

§ 7. Границы применимости классической физики

В дальнейшем появились еще более ошеломляющие экспериментальные факты, свидетельствующие о неприменимости представлений классической физики к явлениям микромира.

Соотношения неопределенностей. Дифракция электронов. В 1927 г. американские физики Л. Джермер и К. Дэвиссон обнаружили, что пучок электронов определенной энергии, прошедший сквозь кристалл и попавший на фотопластинку, давал дифракционную картину такую же, как и прошедший сквозь кристалл пучок рентгеновских или γ-лучей. При этом дифракционная картина не зависела от интенсивности электронного пучка; та же картина получалась в предельном случае весьма слабых пучков, когда можно было считать, что электроны падают на кристалл поодиночке.

Таким образом, оказалось, что электронам, как и фотонам, присущи определенные свойства, характерные для волн, а не для частиц; причем эти волновые свойства приходилось приписывать каждому электрону в отдельности, а не всей совокупности электронов в пучке. Вместе с тем каждый электрон, попадая на фотопластинку, давал почернение только в одном месте, в одном зерне светочувствительного слоя, и лишь совокупность почерневших зерен давала дифракционную картину распределения интенсивности прошедшего пучка. Итак, в одних условиях при прохождении сквозь кристалл электрон вел себя как протяженная волна, а в других — при попадании на зерно фотослоя — как строго локализованная частица. На основании представлений классической физики такое различие в поведении электрона в разных условиях объяснить не удалось.

Зарождение квантовой теории. Все эти явления были правильно объяснены только квантовой механикой. История создания квантовой механики делится на два периода. Первый период — с начала XX столетия и до конца его первой четверти — это период создания так называемой «старой квантовой теории», в основе которой лежат гипотеза Планка о дискретном характере излучения нагретых тел, уравнение Эйнштейна для фотоэффекта и теория атома Бора. Старая квантовая теория не представляла собой стройной, логически замкнутой науки.

Удачно описав некоторые экспериментальные факты, она проявила полную неспособность правильно объяснить и количественно описать все многообразие удивительных явлений микромира.

К концу первой четверти XX века стало ясно, что необходима теория, которая с самого начала выяснила бы, почему классическая физика непригодна при анализе явлений атомного масштаба. Какие же черты классического способа описания физических явлений делают его неприменимым к микрообъектам и где проходит граница применимости представлений классической физики?

Абстракции, лежащие в основе классической физики. Классическое описание физического процесса или явления характеризуется рядом абстракций. Прежде всего это *абсолютизация* понятия физического процесса, заключающаяся в предположении о независимости явлений от условий их наблюдения. Единственное обстоятельство, связанное с условиями наблюдения, которое учитывалось в классической физике, — это выбор системы отсчета: по отношению к двум произвольно движущимся относительно друг друга системам отсчета одно и то же явление будет иметь различный вид. Физический процесс в инерциальной системе отсчета рассматривался как нечто, происходящее независимо от наблюдения за этим процессом, а не как явление, конкретно познаваемое при помощи определенных средств исследования.

Позднейшее развитие физики показало, что абсолютизация физических процессов не является логически необходимой, а представляет собой допущение, которое прекрасно оправдывалось при изучении макроскопических явлений, но которое оказалось совершенно непригодным в микромире.

Действительно, классическая физика имела дело с телами крупного масштаба, по отношению к которым воздействие, связанное с измерением, играло совершенно ничтожную роль. Например, освещение футбольного поля даже самым ярким светом прожекторов и телевизионная съемка матча никак не сказываются на полете футбольного мяча. В тех случаях, когда влияние измерения было заметным (ведь давление света прожекторов в принципе несколько изменяет траекторию мяча), его можно было учесть и внести соответствующие поправки. Принципиальная возможность этого никогда не вызывала сомнений.

Вторая абстракция, допускавшаяся в классической физике, была тесно связана с первой и заключалась в том, что при изучении физических явлений считалось возможной сколь угодно подробная *детализация* описания этих явлений. Другими словами, считалось, что можно неограниченно уточнять наблюдение и наблюдать разные стороны одного и того же физического процесса, не нарушая самого явления. Например, увеличивая частоту кадров телевизионной съемки, можно неограниченно увеличивать точность определения положения и скорости футбольного мяча.

О классическом детерминизме. С этими двумя абстракциями, используемыми в классической физике, с предположением об абсолютном характере физических процессов (в смысле их независимости от условий наблюдения) и о возможности сколь угодно детального их описания (в пределе — исчерпывающе точного и всестороннего) связано понятие о лапласовом механическом детерминизме, согласно которому можно определить состояние исследуемой системы в любой момент времени, коль скоро известно ее начальное состояние.

Вопрос о применимости классического способа описания — это вопрос о возможности использования перечисленных абстракций при анализе конкретного явления. Если в каком-то конкретном случае установлено, что эти абстракции неприменимы, то классическое описание невозможно, и, следовательно, теряют смысл классические представления о свойствах изучаемого объекта, например о его движении по определенной траектории.

Соотношения неопределенностей Гейзенберга. Пределы применимости представлений классической физики, т. е. классического способа описания явлений микромира, устанавливаются так называемыми *соотношениями неопределенностей* Гейзенберга. Рассматривая различные способы измерения положения частицы, Гейзенберг пришел к выводу, что условия, благоприятные для точного измерения положения частицы, неблагоприятны для точного измерения ее импульса и, наоборот, условия, благоприятные для измерения импульса, неблагоприятны для измерения положения частицы. Одно из соотношений Гейзенберга связывает между собой неопределенности в значениях координаты частицы x и соответствующей компоненты импульса p_x в один и тот же момент времени:

$$\Delta x \Delta p_x \geqslant h. \qquad (1)$$

Величины Δx и Δp_x неправильно было бы понимать только как неточности одновременного измерения величин x и p_x, поскольку самый термин «неточность» как бы предполагает, что существуют и «точные» значения x и p_x, но только они почему-то не могут быть измерены. На самом деле по современным представлениям невозможность точного измерения есть следствие того, что частица по своей природе не имеет одновременно точного значения координаты и соответствующей проекции импульса. Эта невозможность есть проявление двойственной корпускулярно-волновой природы материальных микрообъектов. Аналогичные соотношения справедливы и для других координат и компонент импульса:

$$\Delta y \Delta p_y \geqslant h, \qquad (2)$$

$$\Delta z \Delta p_z \geqslant h. \qquad (3)$$

Кроме соотношений (1)—(3) справедливо соотношение, связывающее неопределенность в изменении энергии частицы и неопределенность в моменте времени, когда это изменение произошло:

$$\Delta E \Delta t \geqslant h. \qquad (4)$$

Это соотношение получило название неравенства Бора—Гейзенберга. Оно фактически означает, что определение энергии с точностью до ΔE должно занять промежуток времени, равный по меньшей мере $\Delta t \sim h/\Delta E$. Таким образом, если изучаемая система находится в некотором состоянии в течение времени Δt, то ее энергия имеет неопределенность не менее $\Delta E \sim h/\Delta t$, поскольку Δt — наибольший промежуток времени, в течение которого можно измерять энергию.

Соотношения неопределенностей Гейзенберга являются одним из фундаментальных законов природы. Они справедливы для любых материальных объектов — элементарных частиц, квантов света, атомов, молекул и т. д. Справедливость соотношений неопределенностей, как и всех других фундаментальных законов природы, подтверждается всей совокупностью имеющихся экспериментальных фактов.

Иллюстрация соотношения неопределенностей. Проиллюстрируем неравенство Гейзенберга (1), рассматривая дифракцию плоской световой волны на узкой щели.

Пусть на непрозрачный экран A со щелью шириной Δx падает слева плоская монохроматическая волна (рис. 20). На удаленном экране B (на рис. 20 размер щели сильно преувеличен по сравнению с расстоянием между экранами) наблюдается дифракционная картина, распределение освещенности для которой показано на этом же рисунке.

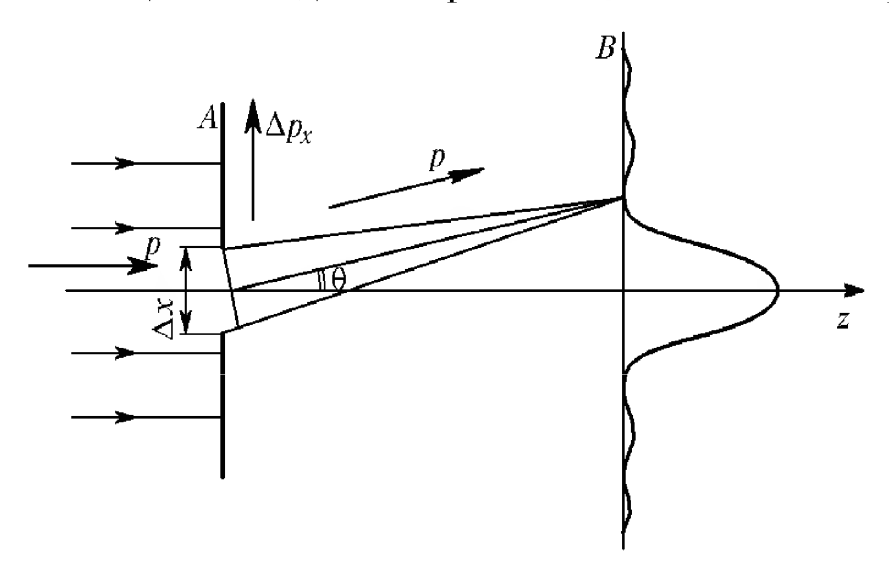


Рис. 20. Дифракция света на щели

Почти весь дифрагировавший свет приходит в область на экране B, ограниченную главным максимумом. Угловую ширину этого максимума легко вычислить. Направление на ближайший минимум характеризуется углом θ, определяемым из условия

$$\Delta x \sin \theta = \lambda. \qquad (5)$$

Рассмотрим теперь эту дифракционную картину с точки зрения представления о свете как о совокупности световых квантов — фотонов. Каждый фотон, прошедший через щель, попадает в определенную точку на экране B. Предсказать, в какую именно точку попадет отдельный фотон, принципиально невозможно. Однако в совокупности большое число попавших на экран B фотонов дает дифракционную картину, представленную на рис. 20. На первый взгляд могло бы показаться, что дифракционную картину можно объяснить взаимодействием между различными фотонами, проходящими через щель, т. е. только в рамках корпускулярных представлений. Однако, уменьшая интенсивность света до таких пределов, когда в любой момент времени между источником света и экраном будет находиться в среднем только один фотон, можно убедиться, что распределение фотонов, попавших на экран за достаточно большой промежуток времени, по-прежнему будет определяться дифракционной картиной. Таким образом, дифракция представляет собой статистическое свойство отдельного фотона.

Проследим, как происходит движение фотона в этом приборе. До щели в экране A распространяется плоская монохроматическая волна, т. е. нам точно известен импульс фотонов

$$p = \frac{h\nu}{c} = \frac{h}{\lambda}, \tag{6}$$

направленный по оси z. Составляющая импульса фотона по оси x равна нулю, т. е. известна точно, но зато совершенно не определена x-координата фотона. При прохождении фотона через щель в экране A ширина щели Δx будет служить мерой неопределенности значения x-координаты фотона. В самом деле, факт появления фотона на экране B позволяет сделать лишь тот вывод, что фотон проник сквозь щель; в какой же именно точке щели это произошло неизвестно. Далее, по корпускулярным представлениям, возникновение на экране дифракционной картины следует истолковать в том смысле, что каждый фотон, пройдя через щель, отклоняется либо вверх, либо вниз. Но для этого фотон должен приобрести составляющую импульса Δp_x, перпендикулярную направлению первоначального движения. Модуль полного импульса фотона p, как видно из формулы (6), при этом не меняется, ибо остается неизменной длина волны.

Поскольку большая часть фотонов попадает в область главного максимума, но принципиально невозможно предсказать, куда попадет каждый фотон, то из рис. 6 ясно, что мера неопределенности x-компоненты импульса Δp_x после прохождения через щель есть

$$\Delta p_x \geqslant p \sin\theta. \tag{7}$$

Перемножая почленно (5) и (7) и учитывая соотношение между импульсом фотона и длиной волны света (6), получаем для момента времени, когда фотон проходит через щель:

$$\Delta x \Delta p_x \geqslant h,$$

что совпадает с формулой (1). Подчеркнем, что проделанный вывод не является доказательством соотношений неопределенностей, а представляет собой лишь иллюстрацию их справедливости для конкретного случая.

Границы применимости классических представлений. Соотношения неопределенностей устанавливают принципиальную границу применимости законов классической физики. Используя их, можно выяснить, справедливы ли представления классической физики для описания конкретного явления. Совершенно очевидно, что для макроскопических объектов — планет, искусственных спутников, артиллерийских снарядов классическое описание является совершенно правильным. Легко убедиться, что при любой достижимой точности измерений координат и импульсов этих объектов соотношения неопределенностей выполняются с огромным запасом и, следовательно, квантовые эффекты никак не проявляются.

Рассмотрим, например, металлический шарик массой 0,01 г. Если мы определим его положение с точностью $\Delta x \approx 0{,}001$ см, доступной нашему зрению в поле микроскопа, то, согласно соотношению неопределенностей, неопределенность скорости такого шарика равна

$$\Delta v = \frac{\Delta p}{m} \approx \frac{h}{m\Delta x} \approx 6\cdot 10^{-22}\ \text{см/с}.$$

Такая точность лежит далеко за пределами возможностей измерений.

Посмотрим, как обстоит дело при изучении свойств более мелких объектов, например электронов. Оказывается, что однозначного ответа на вопрос, применимы ли представления классической физики, в этом случае дать нельзя — все зависит от того, какое явление изучается.

Рассмотрим вначале пучок электронов в кинескопе телевизора. В современном телевизоре ускоряющее напряжение $U \approx 15$ кВ. Разогнанный такой разностью потенциалов электрон обладает импульсом $p = \sqrt{2meU}$. Подставляя в эту формулу значения массы электрона m, его заряда e и ускоряющей разности потенциалов, находим в системе СГС: $p = 6{,}6\cdot 10^{-18}$ г·см/с. Этот импульс направлен вдоль оси трубки. Диаметр пучка, формируемого в современных телевизорах, не бывает меньше $d = 10^{-3}$ см (для телевизора меньший диаметр просто не нужен). Формируя пучок, мы тем самым фиксируем координату электрона в перпендикулярном оси пучка направлении с точностью Δx, равной диаметру пучка d. В силу соотношения неопределенностей при этом электрону сообщается неконтролируемый импульс Δp, перпендикулярный оси пучка:

$$\Delta p \approx \frac{h}{d} \approx 6{,}6\cdot 10^{-24}\ \text{г·см/с}.$$

Связанная с этим неопределенность в направлении движения электрона $\Delta\theta$ определяется отношением

$$\Delta\theta = \frac{\Delta p}{p} \approx 10^{-6} \text{ рад.}$$

Поскольку длина пути электрона в кинескопе не превышает $l \sim 1$ м, то неконтролируемое смещение ΔS электрона на экране, вызываемое квантовыми эффектами, т. е. неопределенностью в направлении движения электрона $\Delta\theta$, не превосходит $\Delta S \leqslant l\Delta\theta = 10^{-4}$ см, т. е. меньше диаметра пучка. Итак, движение электронов в кинескопе телевизора можно рассматривать с помощью законов классической физики.

Рассмотрим теперь электрон в атоме водорода. Известно, что размер атома водорода d равен приблизительно 10^{-8} см. Классическое описание поведения электрона в атоме предполагает, что ему можно приписать определенную траекторию. В планетарной модели атома электрон обращается вокруг ядра. Диаметр его орбиты можно считать равным размеру атома. В классической механике условие его движения по круговой орбите радиуса r — второй закон Ньютона — имеет вид

$$\frac{mv^2}{r} = \frac{e^2}{r^2}.$$

Отсюда получаем для импульса электрона $p = mv$:

$$p = e\sqrt{\frac{m}{r}} \approx 2 \cdot 10^{-20} \ \frac{\text{г}\cdot\text{см}}{\text{с}}. \qquad (8)$$

Однако с помощью соотношения неопределенностей убеждаемся, что если электрон находится внутри атома, т. е. неопределенность в значении его координаты Δx не превосходит размеров атома d, то соответствующая неопределенность в значении импульса Δp оказывается больше, чем сам импульс, вычисляемый по формуле (8):

$$\Delta p \approx \frac{h}{\Delta x} \approx \frac{h}{d} \approx 6{,}6 \cdot 10^{-20} \, \frac{\text{г}\cdot\text{см}}{\text{с}}.$$

Итак, для электрона в атоме классическое описание непригодно. Далее на конкретных примерах будет показано, что с помощью соотношений неопределенностей можно не только убедиться в том, справедливы или нет классические законы в определенных ситуациях, но и исследовать некоторые свойства квантовых объектов.

- Какие опыты свидетельствуют о наличии волновых свойств у электронов?
- Какие абстракции лежат в основе классического способа описания физических явлений?
- В чем заключается абстракция, связанная с абсолютизацией понятия физического процесса? Поясните примерами.

- В чем заключается абстракция, связанная с возможностью сколь угодно подробной детализации описания явлений? Поясните примерами.
- Сформулируйте соотношения неопределенностей Гейзенберга и разъясните их физический смысл.
- Поясните, почему величины Δx и Δp_x в соотношении (1) неправильно понимать только как неточности одновременного измерения координаты x и проекции импульса Δp_x?
- Каким образом явление дифракции света на щели можно использовать для иллюстрации соотношения неопределенностей Гейзенберга для фотонов?
- Поясните, каким образом при дифракции монохроматической волны на щели можно получить условия (5) для направления на ближайший к центральному максимуму минимум дифракционной картины.
- Почему соотношения неопределенностей Гейзенберга не накладывают никаких ограничений на классическое описание движения макроскопических тел?
- Как в конкретных случаях с помощью соотношений неопределенностей можно выяснить возможность использования классической механики для описания движения электрона?

§ 8. Свет — частицы или волны?

Корпускулярно-волновой дуализм. Попробуем использовать соотношения неопределенностей, чтобы разобраться в вопросе: что же все-таки такое свет — частицы или волны? Как мы видели, некоторые оптические явления свидетельствуют в пользу волновых представлений, другие могут быть объяснены только с корпускулярной точки зрения. Наконец, существует целый ряд оптических явлений, которые допускают объяснение как с точки зрения волновых, так и с точки зрения корпускулярных представлений о свете.

Рассмотрев несколько примеров, мы могли убедиться, что при анализе конкретных явлений эта двойственная природа света никак не мешала нашим рассуждениям и не приводила к логическим противоречиям. Нам только нужно было выбрать, на волновом или корпускулярном языке вести рассмотрение, и последовательно придерживаться выбранного способа описания.

Противоречие возникает только тогда, когда мы пытаемся составить общее представление о свете. Действительно, соотношение $E = h\nu$ или $p = h\nu/c$ связывает волновые и корпускулярные свойства фотона: правые части содержат величину ν, определяемую из интерференционных явлений, а левые части, E и p, характеризуют фотон как частицу. Но именно эти-то свойства света и не могут быть логически непротиворечиво объяснены классической физикой, ибо с точки зрения классической физики понятия волны и частицы являются взаимоисключающими.

Частицы и волны в классической физике. Действительно, в классической физике частица всегда движется по определенной траектории и в каждый момент времени имеет определенные координаты и скорость. Если вблизи траектории расположить преграды, которых траектория не задевает, то частица просто «не заметит» таких преград, если, разумеется, они не оказывают на нее силового воздействия посредством каких-либо полей. Для волнового движения характерно понятие протяженного фронта волны (или волновой поверхности), и тем самым отсутствует представление о какой-либо траектории. Помещение преград на пути волны принципиально изменяет условия ее распространения. Различия в свойствах частиц и волн настолько существенны, что не возникает даже мысли о возможности сочетания этих свойств у одного объекта. И тем не менее, в явлениях микромира оказалось, что эти свойства проявляются у одних и тех же объектов.

Мысленный опыт, снимающий логические противоречия. Для иллюстрации возникающих логических трудностей, а также для демонстрации того, как они преодолеваются квантовой теорией, рассмотрим подробнее уже упоминавшийся выше простой мысленный дифракционный опыт, схематически представленный на рис. 21. Реальные опыты, которые могут быть осуществлены, более

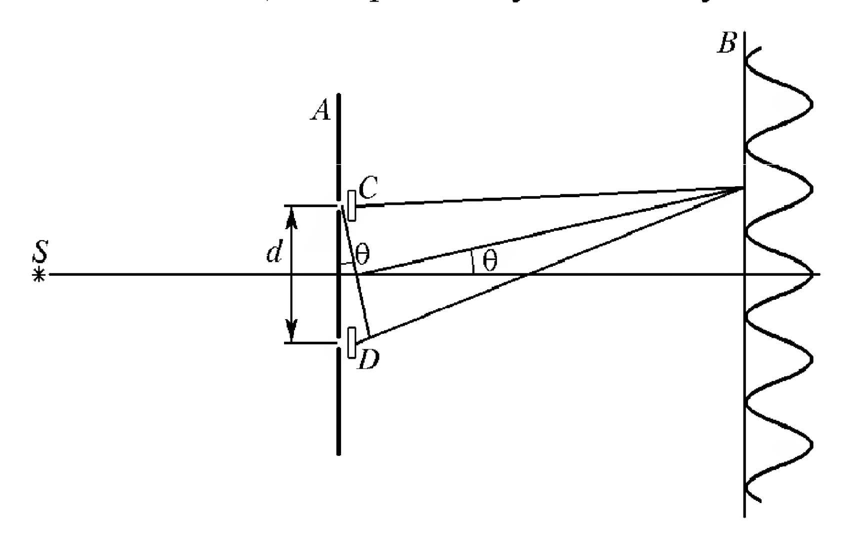


Рис. 21. Дифракция света на двух щелях

сложны, но их результаты подтверждают справедливость заключений, сделанных в результате анализа обсуждаемого здесь мысленного опыта.

Источник света S освещает экран A, в котором прорезаны две щели. Расстояние от A до B велико по сравнению с расстоянием d между щелями, которое в свою очередь много больше длины световой волны. На светочувствительном экране B возникает дифракционная картина, причем в местах дифракционных максимумов вырывается наибольшее число фотоэлектронов. Как и в разобранном выше опыте с дифракцией на одной щели, эксперимент показывает, что дифракционная картина сохранится и в том случае, если в каждый момент вре-

мени между источником и экраном в среднем будет находиться только один фотон. Распределение множества фотонов, попавших на экран за достаточно большой промежуток времени, по-прежнему будет определяться классической картиной дифракции от двух щелей, хотя при вырывании фотоэлектронов из экрана *B* фотоны ведут себя как частицы, каждая из которых выбивает электрон в определенном месте экрана.

Если закрыть одну из щелей, то интерференционные полосы пропадают — распределение интенсивности на экране становится таким же, как при дифракции на одной щели, и при очень узкой щели становится практически равномерным. Поэтому мы вынуждены считать, что при движении от источника света через щели до экрана *B* излучение ведет себя как волна. Если попытаться объяснить результаты опыта с помощью представления о свете как о частицах, то нужно считать, что каждый фотон, по-видимому, проходит только через одну из щелей. Но тогда, в рамках чисто корпускулярных представлений, можно было бы спросить: каким образом поток независимых фотонов, каждый из которых проходит только через одну из щелей, может образовать дифракционную картину, наблюдаемую лишь при наличии обеих щелей? Или, другими словами, каким образом щель, через которую фотон не проходит, не позволяет ему попасть на те места экрана, куда он мог бы попасть, если бы эта щель была закрыта?

В этой формулировке вопроса предполагается, что фотон действительно проходит через одну из щелей. С точки зрения классической теории это допущение является естественным, ибо предполагается, что в любой момент времени фотон (как и любая другая частица) движется по определенной траектории, т. е. имеет определенные координаты, доступные измерению. Современная квантовая теория отказывается от этого допущения, утверждая, что говорить о положении фотона имеет смысл лишь в том случае, если при постановке опыта мы позаботимся об определении его координаты. Значит, если мы хотим считать, что каждый фотон действительно, подобно частице, проходит только через одну из щелей, мы должны поставить какой-либо измерительный прибор, который бы фиксировал нам факт прохождения фотона через определенную щель. Если мы попробуем с помощью специальных счетчиков *C* и *D* фиксировать, через какое отверстие проходит каждый фотон, то обнаружим, что дифракционная картина на экране *B* размоется. Попробуем объяснить этот экспериментальный факт, используя соотношения неопределенностей Гейзенберга. Выясним, можно ли в принципе наблюдать на экране *B* интерференционные полосы, если точно определять, через какие отверстия проходит фотон.

Если с помощью счетчиков, установленных непосредственно вблизи отверстий в экране *A*, мы будем определять, через какое именно отверстие проходит каждый фотон, то тем самым мы

действительно заставим свет проявлять корпускулярные свойства, ибо только для частицы имеет смысл утверждение, что она прошла через определенное отверстие, для волны вопрос о том, через какое отверстие она прошла (разумеется, если открыты оба), вообще лишен смысла.

Для того чтобы уверенно судить о том, через какое отверстие прошел фотон, нужно с помощью счетчиков определять его координату x с достаточной точностью, так чтобы ошибка Δx в определении координаты фотона была бы меньше половины расстояния между отверстиями:

$$\Delta x < \frac{d}{2}. \tag{1}$$

Однако не следует стремиться определять x-координату фотона слишком точно, так как в силу соотношения неопределенностей Гейзенберга

$$\Delta x \Delta p_x \geqslant h \tag{2}$$

это приводит к слишком большой неопределенности в значении x-компоненты импульса фотона и дифракционная картина на экране B заведомо размажется, т. е. нам тогда не удастся наблюдать в этом опыте проявление волновых свойств света.

Итак, для наблюдения одновременного проявления светом корпускулярных и волновых свойств необходимо определять координату фотона с максимальной неопределенностью, совместимой с выражением (1), т. е. с возможностью установления, через какое отверстие прошел фотон, надеясь при этом, что вносимая при измерении координаты неопределенность в значении импульса Δp_x еще не приведет к полному размытию интерференционных полос. Выясним, совместимы ли эти условия.

Какова максимально допустимая неопределенность в значении импульса Δp_x, которая еще не приводит к полному уничтожению дифракционной картины? Дифракционная картина от двух отверстий, наблюдаемая на экране B, состоит из чередующихся светлых и темных полос. Угловое расстояние $\Delta\theta$ между направлениями на соседние максимум и минимум определяется из условия $d\Delta\theta = \lambda/2$, поскольку максимумы расположены в тех точках экрана B, разность хода до которых от отверстий в экране A равна целому числу длин волн. Неопределенность в значении импульса Δp_x можно выразить через неопределенность направления импульса $\Delta\theta_1$: $\Delta p_x = p\Delta\theta_1$. Интерференционные полосы, очевидно, не будут полностью размыты, только если $\Delta\theta_1 < \Delta\theta$, т. е.

$$\frac{\Delta p_x}{p} < \frac{\lambda}{2d}.$$

Используя соотношение между импульсом фотона и длиной волны $\lambda = h/p$, можно переписать это неравенство в виде

$$2d\Delta p_x < h. \qquad (3)$$

Таким образом, для проявления волновых свойств света в этом опыте должно выполняться неравенство (3), а для проявления корпускулярных свойств — неравенство (1). Объединяя эти неравенства, получаем условие одновременного проявления светом корпускулярных и волновых свойств:

$$\Delta x \Delta p_x < h/4.$$

Но это условие противоречит соотношению неопределенностей, а потому не может быть выполнено. Итак, установив, через какие отверстия проходят фотоны, мы теряем дифракционную картину и не можем говорить о проявлении фотонами волновых свойств. Мы рассмотрели упрощенный мысленный опыт. В более сложном варианте подобные опыты действительно были поставлены и полностью подтвердили результаты приведенного здесь анализа.

Подведем некоторые итоги. Как мы видели, отдельные фотоны обнаруживают волновое поведение, заключающееся в том, что каждый фотон проявляет интерференционные свойства независимо от других фотонов. Но, попадая на экран, фотоны обнаруживают корпускулярное поведение, заключающееся в том, что они взаимодействуют с веществом только в определенных точках. Если при этом не делать попытки экспериментально наблюдать траектории фотонов до попадания их на экран, то, пропустив большое число фотонов, мы будем наблюдать на экране дифракционную картину, предсказываемую волновой теорией. Но предсказать, в какое место экрана попадет определенный фотон, невозможно. Это можно сделать только в вероятностном смысле: вероятность попасть фотону в область минимума мала. Для проверки такого предсказания нужно большое число фотонов. Если же фиксировать траектории фотонов до попадания на экран, то фотоны вовсе не проявляют волновых свойств.

Фотон — квантовый объект. Как же понимать тот факт, что фотоны выступают иногда в облике частиц, а иногда в облике волн? Квантовая теория отвечает на этот вопрос так: фотон представляет собой квантовый объект, а когда мы описываем его поведение как поведение частицы или волны, мы навязываем классическое описание этому объекту, имеющему существенно неклассическую природу. Рассматривать поведение фотона имеет смысл, только исходя из результатов измерений, совершаемых над ним. Поэтому то, как поведет себя фотон — как частица или как волна, зависит от характера проводимого над ним измерения.

Итак, что же такое свет — частицы или волны? Ни то, ни другое. Мы можем заставить материальный объект, который мы назы-

ваем светом, проявлять либо корпускулярные, либо волновые свойства. Но в принципе невозможно осуществить эксперимент, в котором свет одновременно проявлял бы и те, и другие свойства. Этим устраняется логическая трудность. Действительно, нам больше не нужно пытаться представить себе, как это фотон может быть одновременно и волной, и частицей. Теперь корпускулярно-волновой дуализм мы понимаем в том смысле, что свет обладает *потенциальной возможностью* проявлять и волновые, и корпускулярные свойства, но они никогда не проявляются одновременно. Соотношения $E = h\nu$ и $p = h\nu/c$ означают, что между этими взаимно исключающими друг друга свойствами имеет место эквивалентность в том смысле, что меры этих свойств всегда пропорциональны. Эти свойства *дополняют* друг друга, ибо только их совокупность дает полное представление о свете. Но, как показывают соотношения неопределенностей, в любом явлении в зависимости от конкретных условий реализуется только одна возможность.

Корпускулярно-волновой дуализм присущ не только фотонам, но и любым другим микрообъектам.

Волны де Бройля. Мы уже упоминали об экспериментах, в которых наблюдалась дифракция электронов при прохождении их сквозь кристаллы. Но еще до осуществления таких экспериментов, в 1924 г., Луи де Бройль предположил, что все частицы должны обладать волновыми свойствами, подобными волновым свойствам света, и ввел количественное соотношение между длиной сопоставляемой частице волны и импульсом частицы, аналогичное соотношению между длиной волны и импульсом фотона:

$$\lambda = \frac{h}{p}. \tag{4}$$

Несколько лет спустя Джермер и Дэвиссон, изучая рассеяние электронов кристаллами, обнаружили дифракцию электронов, подобную дифракции света на решетке. Атомы кристалла никеля, который использовался в опыте, образуют регулярную конфигурацию, которая действует подобно дифракционной решетке. Максимумы в распределении рассеянных кристаллом электронов находились на тех местах, для которых выполнялось условие

$$n\lambda = d \sin \theta. \tag{5}$$

Но это условие совпадает с известным условием для максимумов при дифракции волн на решетке с периодом d. Вычисленное по (5) значение длины волны λ совпадало со значением, даваемым (4), с погрешностью до 1%. Следовательно, электроны, отражаясь от кристалла, дифрагируют точно так же, как если бы они были волнами с длиной, предсказанной де Бройлем.

Волновые свойства микрочастиц. Результаты этого опыта имеют фундаментальное значение, ибо они демонстрируют волновые свойства вещества, которые не могут быть поняты в рамках представлений о том, что вещество состоит из классических частиц. Более поздние опыты показали, что и другие частицы вещества, даже такие относительно крупные, как молекулы, также проявляют волновые свойства.

Когда реальность проявления волновых свойств микрочастицами стала очевидной, возникла необходимость как-то интерпретировать волны де Бройля, придать им определенный физический смысл. Появилась концепция «волны-пилота», в которой предполагалось, что волна в каком-то смысле «управляет» движением частицы. Но целый ряд экспериментальных фактов показывал, что такое весьма наглядное представление не приводит к внутренне непротиворечивой картине поведения микрочастиц. Правильное толкование этого вопроса стало возможным только после создания квантовой механики. Пока мы будем просто считать, что соотношение (4) дает нам длину волны, которую следует сопоставлять любому материальному объекту, если опыт показывает проявление этим объектом волновых свойств.

Еще о границах классического описания. Представления о волнах де Бройля можно, наряду с соотношениями неопределенностей Гейзенберга, использовать для выяснения вопроса о том, какой теорией, квантовой или классической, следует описывать конкретное явление. Для этого нужно сравнить сопоставляемую по формуле (4) изучаемому объекту длину волны с характерными размерами в соответствующей задаче, имея в виду, что волновые свойства объекта не играют существенной роли, пока эта длина волны не станет соизмеримой с характерными размерами. Например, длина волны электрона, находящегося в атоме водорода, сравнима с размерами самого атома. Поэтому квантовые эффекты в этом случае будут весьма существенными, и представления классической физики здесь заведомо неприменимы.

Сравним теперь длину волны де Бройля, сопоставляемую Земле, с размерами земной орбиты. Так как масса Земли $M = 6 \cdot 10^{27}$ г, скорость Земли на орбите $v \approx$ 30 км/с, а расстояние от Земли до Солнца $R \approx 15 \cdot 10^7$ км, то

$$\frac{\lambda}{2\pi R} = \frac{h}{Mv \cdot 2\pi R} \approx 3 \cdot 10^{-75}.$$

Эта величина фантастически мала. Следовательно, движение Земли будет превосходно описываться классической механикой. Любые волновые или квантовые эффекты будут в этом случае меньше, чем, например, эффекты, вызванные столкновением Земли с протоном или электроном, содержащимся в космических лучах.

Задача

Для уменьшения размеров пятна на экране электронно-лучевой трубки можно на пути разогнанного ускоряющим напряжением электронного пучка поместить две диафрагмы с отверстиями (рис. 22). Покажите, что для второго отверстия существует оптимальный диаметр, обеспечивающий наименьший размер пятна на экране.

Решение. После прохождения ускоряющего промежутка с напряжением U электроны в трубке движутся равномерно с практически одинаковыми по модулю импульсами p, определяемыми из соотношения

$$\frac{p^2}{2m} = eU. \tag{6}$$

Будем сначала рассматривать электрон как классическую частицу. Пусть отверстие в первой диафрагме настолько мало, что его можно считать точечным. В этом случае диаметр пятна на экране был бы тем меньше, чем меньше отверстие во второй диафрагме. Угловой размер пятна, как видно из рис. 23, определяется соотношением

$$\theta = d/l, \tag{7}$$

где d — диаметр отверстия во второй диафрагме, а l — расстояние между диафрагмами. Таким образом, пятно от классических частиц можно было бы сделать сколь угодно малым. Однако из-за квантовых явлений уменьшение диа-

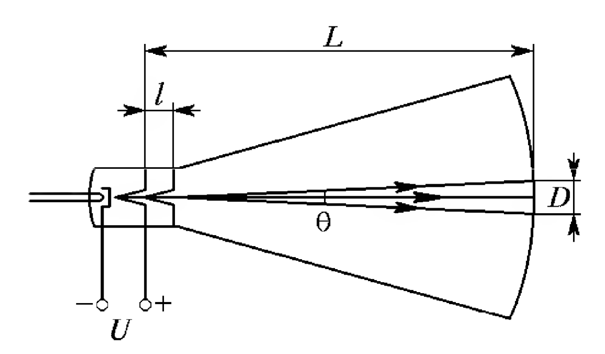


Рис. 22. Для уменьшения размера пятна на экране можно использовать две диафрагмы с отверстиями

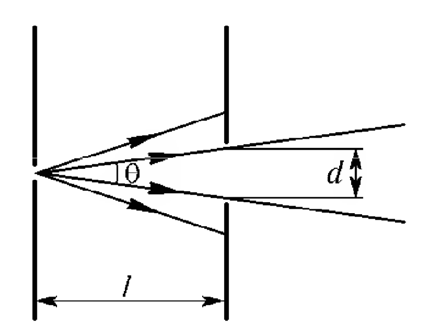


Рис. 23. Угловой размер пучка θ зависит от диаметра d отверстия во второй диафрагме

метра диафрагмы с некоторого момента приводит к расширению пятна. Оценить критический диаметр отверстия можно с помощью соотношения неопределенностей Гейзенберга.

Если электрон прошел сквозь отверстие во второй диафрагме, то неопределенность в значении его координаты в направлении поперек пучка Δx определяется размером отверстия d:

$$\Delta x \approx d. \tag{8}$$

Такая локализация электрона приводит к появлению у него неконтролируемой составляющей импульса поперек пучка:

$$\Delta p_x \approx h/\Delta x \approx h/d. \tag{9}$$

В результате после прохождения диафрагмы появляется неопределенность в направлении движения электрона, характеризуемая углом $\theta_{\text{кв}}$:

$$\theta_{\text{кв}} = \frac{\Delta p_x}{p} \approx \frac{h}{pd}. \tag{10}$$

Из (10) следует, что при уменьшении диафрагмы отверстия d происходит расширение пучка и как следствие увеличение пятна на экране трубки. Уменьшать отверстие в диафрагме следует до тех пор, пока размытие пучка $\theta_{кв}$ не сравняется с его угловым размером θ, определяемым классическими траекториями электронов: $\theta_{кв} = \theta$. Отсюда с помощью (7) и (10) определяем оптимальный размер отверстия во второй диафрагме:

$$d \approx \sqrt{hl/p}. \tag{11}$$

Этот результат можно получить иначе, используя представление о связанной с электроном волне де Бройля

$$\lambda = h/p. \tag{12}$$

Если классическому представлению о движении электронов по определенным траекториям соответствует приближение геометрической оптики, то проявлению квантовых свойств отвечает дифракция волн де Бройля. Дифракционные явления дают отклонение от геометрического закона распространения лучей на углы порядка отношения длины волны к размеру препятствия или отверстия. Поэтому

$$\theta_{волн} = \lambda/d.$$

Подставляя сюда вместо λ де-бройлевское выражение (12) для длины волны электрона, приходим к прежнему соотношению (10).

- Оцените оптимальный диаметр отверстия в трубке с такими характеристиками: ускоряющее напряжение $U = 10$ кВ, расстояние между диафрагмами $l = 1$ см, расстояние до экрана $L = 0{,}5$ м.

- Поясните, в каком смысле в соотношениях $E = h\nu$ и $p = h\nu/c$ левые части характеризуют фотон как частицу, а правые — как волну.
- Почему в классической физике представления о волне и о частице неприменимы к какому бы то ни было одному и тому же объекту?
- Поясните, каким образом в мысленном опыте по дифракции на двух щелях происходит полное размытие интерференционных полос, если фиксировать щель, через которую проходит каждый из фотонов.
- Покажите, почему возможность проявления электроном взаимоисключающих корпускулярных и волновых свойств не приводит к логическому противоречию.
- Что такое волны де Бройля? Как длина волны де Бройля зависит от массы и скорости частицы?
- Как применить представление о волнах де Бройля к выяснению возможности классического описания?

§ 9. Законы движения в квантовой физике

Итак, мы видели, что многие явления в микромире не описываются классической физикой и, пользуясь соотношениями неопределенностей Гейзенберга, можно установить границы применимости классического способа описания при рассмотрении тех или иных конкретных явлений. В тех случаях, когда классическое описание оказывается не-

пригодным, необходим более совершенный способ описания физических явлений, который должен учитывать возможность проявления изучаемыми объектами как корпускулярных, так и волновых свойств.

Современная квантовая теория ведет свое начало с 1926 г., когда Э. Шрёдингером было предложено уравнение, носящее ныне его имя и лежащее в основе квантовой механики. Разумеется, изложение квантовой механики выходит за рамки школьного курса физики, но мы можем обсудить разобранные выше экспериментальные факты и четко сформулировать, какой должна быть квантовая теория, способная последовательно объяснить все своеобразие явлений микромира.

Роль средств наблюдения в квантовой физике. Обсуждая причины неприменимости представлений классической физики в микромире, мы видели, что эта неприменимость обусловлена рядом абстракций, допускавшихся в классической физике. В классической физике молчаливо предполагалась независимость физических процессов от способов наблюдения и возможность наблюдать одновременно все стороны данного процесса. В области квантовых явлений это не так. Вспомните, например, опыт с дифракцией фотонов на двух щелях: определяя, через какое отверстие проходит каждый фотон, мы этим измерением принципиально изменяли протекание физического процесса, так что дифракционная картина на экране оказывалась полностью размытой.

Анализируя этот опыт, мы приходим к выводу, что основой нового способа описания явлений должен быть явный учет реальных *возможностей измерений*, проводимых над микрообъектами. Необходимым посредником при изучении таких объектов являются *приборы*: атомный объект может проявить свои свойства, только провзаимодействовав с прибором. Например, путь микрочастицы становится видимым в результате конденсации паров в камере Вильсона или в результате почернения зерен фотоэмульсии, и т. п. При этом приборы и условия опыта должны описываться *классически*, путем задания значений параметров, характеризующих приборы. Параметры этих приборов могут, разумеется, задаваться лишь с точностью, допускаемой соотношениями неопределенностей.

Вероятность в классической и квантовой физике. В основу нового способа описания поведения микрообъекта следует положить результаты взаимодействия этого объекта с классически описываемым прибором. Свойства атомного объекта выводятся из рассмотрения результатов таких взаимодействий. Это не исключает возможности введения таких величин, которые характеризуют сам микрообъект независимо от прибора (заряд, масса частицы и т. д.), но в то же время позволяет изучать поведение объекта с той его стороны (например, корпускулярной или волновой), проявление которой обусловлено устройством прибора. Таким образом, появляется возможность рассматривать и тот случай, когда разные стороны и разные

свойства объекта не проявляются одновременно. По Бору, свойства, проявляющиеся при *взаимно исключающих* условиях (вспомните обсуждение корпускулярно-волнового дуализма), *дополняют* друг друга в том смысле, что только их совокупность характеризует объект полностью. В этом заключается сформулированный Н. Бором *принцип дополнительности*, который сыграл огромную роль в обосновании современной квантовой теории. Рассматривать одновременное проявление дополнительных свойств не имеет смысла. Таким образом, в новом подходе и не возникает внутреннего противоречия в понятии «корпускулярно-волновой дуализм».

Итак, в основе описания — результаты взаимодействия микрообъекта с прибором. Но опыт показывает, что при данных внешних условиях результат взаимодействия объекта с прибором не является однозначно определенным, а обладает лишь некоторой вероятностью. Вспомните тот же дифракционный опыт: каждый фотон попадал в определенное место экрана, но предсказать точно, в какое именно, было невозможно; существовала лишь определенная вероятность попадания фотона в то или иное место. Таким образом, в описание микрообъекта, его состояния и поведения вводится новый элемент — понятие *вероятности*, а тем самым и понятие *потенциальной возможности.*

Понятие вероятности используется и в классической физике при изучении свойств систем, состоящих из большого числа частиц. Вероятности вводятся тогда, когда условия опыта не известны полностью и по неизвестным параметрам приходилось проводить усреднение. Например, при рассмотрении движения молекул в газе нам неизвестны координаты и скорости каждой молекулы. Поэтому можно предсказать только вероятность попадания частицы в то или иное место. В классической физике вероятности отражают неполноту формулировки задачи, которая, быть может, практически и неизбежна, но в принципе устранима.

Описание состояния системы в квантовой физике. В квантовой физике вероятности имеют совсем иной характер. Здесь они принципиально необходимы; их введение характеризует не полноту условий, а объективно существующие при данных условиях потенциальные возможности. Следует отметить, что в процессе создания квантовой теории высказывалась точка зрения, что введение понятия вероятности и в квантовой механике все-таки связано с тем, что на самом деле микрообъекты обладают определенной, не известной нам внутренней структурой. Поэтому в квантовой механике, как и в статистической физике, приходится считать, что переменные, описывающие эту внутреннюю структуру, распределены случайным образом. Иными словами, в квантовую теорию понятие вероятности вводится только потому, что она не является полной; она станет полной только тогда, когда будут найдены величины, характеризующие внутреннюю структуру частиц. Но в настоящее время можно

утверждать, что квантовая теория, основанная на понятии вероятности результата взаимодействия объекта с прибором, оказалась исключительно успешной и, наоборот, нет никаких экспериментальных фактов, которые свидетельствовали бы о неполноте квантовомеханического описания поведения микрообъектов.

Продолжим обсуждение основ квантовой теории. Из сказанного выше ясно, что задать состояние квантового объекта — значит задать распределение вероятностей, или потенциальных возможностей получения того или иного результата взаимодействия объекта с прибором. То обстоятельство, что должны задаваться вероятности также и для тех величин, измерения которых несовместимы, показывает, что речь идет именно о потенциальных возможностях, а не о значениях величин самих по себе, вне связи с условиями их измерения на опыте.

Принцип соответствия. Теперь можно сказать несколько слов и о том, каким должен быть математический аппарат квантовой теории по сравнению с математическим аппаратом классической физики. В классической физике математический аппарат должен давать значения определенных физических величин, т. е. числа. В квантовой теории математический аппарат должен давать не только возможные значения физических величин, но и вероятности получения на опыте тех или иных возможных значений этих величин.

В квантовой теории сформулированные требования могут быть выполнены следующим образом. Состояние изучаемой системы характеризуется определенной функцией, называемой волновой или ψ-функцией. Волновая функция характеризует состояние системы в том смысле, что через эту функцию выражаются все вероятности для результатов измерения над системой. Волновая функция определяется из уравнения Шрёдингера — основного уравнения квантовой механики. Каждой физической величине сопоставляются определенные математические операции, которые следует проделать над ψ-функцией, чтобы получить необходимую информацию. Отсюда возникает понятие *оператора* соответствующей физической величины. В квантовой механике каждой физической величине, например энергии, координате, импульсу, сопоставлен определенный оператор.

Как выбираются сами операторы? При построении квантовой теории огромную роль сыграл так называемый принцип соответствия, сформулированный Бором: законы квантовой физики должны быть сформулированы таким образом, чтобы в классических границах, когда, например, в изучаемый процесс вовлечено много квантов, эти законы приводили бы к классическим уравнениям для усредненных величин. Использование принципа соответствия позволило найти вид операторов, сопоставляемых определенным физическим величинам. Но общих правил составления операторов для физических величин указать нельзя.

Требование удовлетворения принципу соответствия отнюдь не является тривиальным. Возникает вопрос, как согласовать кванто-

вый подход, основанный на рассмотрении вероятностей, с классическим, допускающим точное предсказание поведения системы. Рассмотрим с этой точки зрения уже разобранный выше пример рассеяния фотонов свободными электронами. Как видно из формулы (11) § 6 «Световые кванты», увеличение длины волны при единичном акте рассеяния фотона равно $2h/(m_0c)$, что примерно составляет 10^{-10} см. Это слишком маленькая величина, чтобы ее можно было заметить при рассеянии радиоволн с длиной волны порядка 1 см и больше. Соответствующие таким волнам частоты оказываются порядка 10^{10} Гц, так что энергия одного фотона составляет 10^{-16} эрг, или 10^{-4} эВ. Легко подсчитать, какую наибольшую энергию может приобрести электрон в результате рассеяния одного такого фотона. С помощью формулы (10) § 6 находим, что $\Delta E_{max} = h\Delta\nu_{max}$ составляет 10^{-14} эВ. Для того чтобы в результате рассеяния радиоволн приобрести энергию всего в 1 эВ, электрон должен рассеять по меньшей мере 10^{14} квантов! Разумеется, нельзя точно предсказать результат каждого индивидуального акта рассеяния. Но результаты рассеяния такого большого числа квантов, которое фактически представляет собой непрерывный процесс, практически вполне определенные и совпадают с тем, что дает для этого случая классическая электродинамика.

Мы построили качественную картину квантовой механики и пояснили на примере, что в пределе, при переходе к классическим условиям, результаты квантовой теории переходят в результаты, даваемые классической физикой. С помощью соотношений неопределенностей можно делать оценки границы применимости законов классической физики. Однако эта граница не является четко очерченной, и развитие физики не раз давало примеры вторжения квантовых эффектов в область, кажущуюся абсолютно классической. С некоторыми из этих примеров мы еще встретимся ниже.

- Объясните, чем обусловлена особая роль средств наблюдения в квантовой физике по сравнению с классической.
- Разъясните содержание принципа дополнительности Бора. Каким образом этот принцип снимает противоречия в понятии «корпускулярно-волновой дуализм»?
- В чем принципиальное различие использования понятия вероятности в классической и в квантовой физике?
- Как задается состояние физической системы в квантовой физике?
- Чем отличаются требования к математическому аппарату в классической и в квантовой физике?
- Что такое волновая функция и что такое оператор физической величины?
- В чем заключается принцип соответствия Бора? Поясните его роль в физике.

III. АТОМЫ. МОЛЕКУЛЫ. КРИСТАЛЛЫ

§ 10. Строение атома

После открытия Дж. Томсоном электронов стало ясно, что эта частица входит важнейшей составной частью во все атомы. Основанная на этой гипотезе исторически первая модель атома была предложена самим Дж. Томсоном.

Томсоновская модель атома. Модель Томсона была построена так, чтобы можно было объяснить наблюдаемое на опыте испускание возбужденными невзаимодействующими между собой атомами почти монохроматического света.

Для испускания монохроматического света электрон в возбужденном атоме, по классическим представлениям, должен совершать гармоническое колебание соответствующей частоты. Поэтому при смещении электрона в атоме из положения равновесия на него должна действовать квазиупругая, т. е. пропорциональная смещению, возвращающая сила. Так получится, если считать, что атом представляет собой шар, положительный заряд которого, равный модулю полного отрицательного заряда электронов, равномерно размазан по всему объему атома.

Томсоновской модели атома была уготована очень недолгая жизнь, хотя отдельные ее положения, в частности представление о квазиупругой силе, удерживающей электрон в атоме, сохранили свое значение до настоящего времени. Это представление используется в теории взаимодействия света с веществом, объясняющей дисперсию света, т. е. зависимость скорости света в веществе от длины волны.

Опыты Резерфорда. В начале девятисотых годов Резерфорд приступил к своим знаменитым опытам по зондированию атомов с помощью альфа-частиц, испускаемых радиоактивными веществами. Масса этих частиц примерно в 8000 раз больше массы электрона, положительный заряд равен удвоенному элементарному электрическому заряду. Скорость испускаемых радиоактивным препаратом альфа-частиц составляет около 1/15 скорости света. Такими частицами Резерфорд бомбардировал атомы тяжелых элементов.

Электроны вследствие своей малой массы не могут заметно изменить траекторию альфа-частицы. Изменение траектории, т. е.

рассеяние альфа-частиц, может вызываться только тяжелой положительно заряженной частью атома. Поэтому из опытов по рассеянию альфа-частиц можно определить характер распределения положительного заряда и массы внутри атома.

Схема опытов Резерфорда показана на рис. 24. Из испускаемых радиоактивным изотопом альфа-частиц с помощью диафрагм выделяется пучок, который падает на тонкую фольгу из исследуемого материала. Рассеянные альфа-частицы попадают на экран, покрытый сернистым цинком. Попадание каждой альфа-частицы создает короткую вспышку света (сцинтилляцию), которую можно наблюдать в микроскоп.

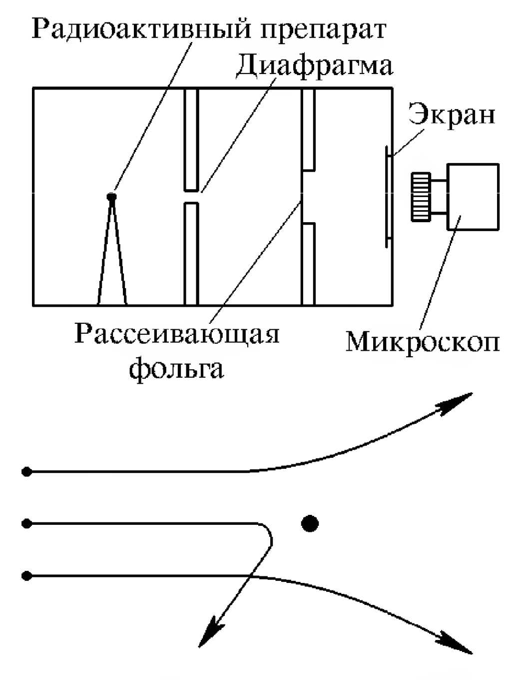


Рис. 24. Схема опытов Резерфорда

Рассеяние на большие углы. При изучении распределения рассеянных альфа-частиц было обнаружено, что некоторые из них отклонились от первоначального направления на очень большие углы. Этот результат был очень неожиданным. «Это почти столь же невероятно, говорил Резерфорд, — как если бы вы выстрелили 15-дюймовым снарядом в кусок тонкой бумаги, а снаряд возвратился бы к вам и нанес вам удар». Действительно, если считать электрический заряд атома распределенным по всему его объему, то создаваемое им электрическое поле будет настолько слабым, что не сможет отклонить пролетающую альфа-частицу на заметные углы.

Открытие атомного ядра. Очевидно, что альфа-частица может быть отброшена назад лишь при условии, что положительный заряд атома и его масса сосредоточены в очень малом объеме внутри атома. Таким образом было открыто *атомное ядро* — тело малых размеров, в котором сконцентрированы весь положительный заряд атома и почти вся его масса.

Изучая подсчетом вспышек в разных местах экрана распределение рассеянных альфа-частиц по углам, Резерфорд смог оценить размеры ядра. Оказалось, что ядра имеют диаметр порядка 10^{-13}–10^{-12} см. Это в сотни тысяч раз меньше диаметра самого атома. Из этих же опытов удалось впоследствии определить и заряд ядра. Если принять элементарный электрический заряд за единицу заряда, то заряд любого ядра в точности равен номеру данного химического элемента в таблице Менделеева.

Планетарная модель атома. Опыты Резерфорда естественным путем приводят к *планетарной модели атома.* В центре атома находится положительно заряженное ядро, в котором сосредоточена почти вся масса атома. Так как атом в целом электрически нейтрален, то число электронов в атоме, как и заряд его ядра (в единицах элементарного заряда), равно порядковому номеру элемента в периодической системе. Электроны благодаря кулоновскому притяжению к ядру обращаются вокруг него подобно тому, как планеты обращаются вокруг Солнца.

Планетарная модель атома представляется совершенно необходимой для объяснения опытов по рассеянию альфа-частиц. Однако она не способна объяснить сам факт сколько-нибудь продолжительного существования атома, т. е. его устойчивость. В самом деле, согласно законам классической электродинамики движущийся с ускорением заряд должен излучать уносящие энергию электромагнитные волны. Поэтому электроны должны были бы за очень короткое время ($\sim 10^{-8}$ с) растерять всю свою энергию и упасть на ядро. Однако в действительности невозбужденные атомы устойчивы и могут существовать неограниченно долго, не излучая электромагнитных волн. Опять налицо явное противоречие наблюдаемых на опыте свойств с представлениями классической физики.

Постулаты Бора. Исторически первая попытка разрешения этого противоречия была предпринята Н. Бором, который сформулировал постулаты, несовместимые с классической механикой и электродинамикой, но позволившие многое понять в поведении атомных систем.

Первый постулат Бора гласит: атомная система может находиться только в особых *стационарных*, или квантовых, состояниях, каждому из которых соответствует определенная энергия системы. Поэтому возможные значения энергии атома образуют дискретный набор. В стационарном состоянии атом не излучает.

Согласно второму постулату Бора излучение света происходит при переходе атома из квантового состояния с большей энергией E_n, в квантовое состояние с меньшей энергией E_m. Энергия $h\nu$ излученного фотона равна разности энергий квантовых состояний:

$$h\nu = E_n - E_m. \qquad (1)$$

Возможен и обратный процесс, в котором атом переходит из квантового состояния с меньшей энергией в состояние с большей энергией. При этом происходит поглощение фотона с энергией, равной разности энергий этих квантовых состояний. Для элементарных актов перехода атома из одного квантового состояния в другое с испусканием или поглощением фотона выполняется закон сохранения энергии.

На основе этих постулатов Бору удалось построить теорию простейшего атома — атома водорода, содержащего только один электрон. Эта теория позволила объяснить установленные на опыте швейцарским учителем физики И. Бальмером закономерности в линейчатом спектре излучения атома водорода.

Правило квантования. Для определения стационарных состояний атома водорода Бор в 1913 г. предложил определенное «правило квантования», согласно которому момент импульса электрона, обращающегося вокруг ядра, может принимать только дискретные значения, кратные постоянной Планка:

$$\hbar = h/2\pi. \tag{2}$$

Правило квантования, приводящее к правильным, согласующимся с опытом значениям энергий стационарных состояний атома водорода, было по сути дела просто угадано Бором. Де Бройль смог дать физическую интерпретацию правилу квантования, основываясь на представлениях о волновых свойствах электрона. Он предположил, что каждая допустимая орбита электрона в атоме водорода соответствует волне, распространяющейся по кругу около ядра атома. Стационарная волна, отвечающая электрону в разрешенном квантовом состоянии, может быть получена, только если волны непрерывно повторяют себя после каждого полного оборота вокруг ядра.

Здесь можно увидеть аналогию со стационарной картиной стоячих волн в струне с закрепленными концами. В этих примерах из классической физики даже для непрерывных волн тоже возникает дискретный набор разрешенных частот. Таким образом, появление дискретных значений физических величин в квантовой физике связано с волновыми свойствами электронов, входящих в атомную систему.

Итак, в стационарном квантовом состоянии атома водорода на длине орбиты электрона должно укладываться целое число длин волн де Бройля λ, т. е.

$$n\lambda = 2\pi r, \tag{3}$$

где r — радиус круговой орбиты. Подставляя в это соотношение длину волны де Бройля $\lambda = h/p$, получим

$$n\,\frac{h}{p} = 2\pi r,$$

или

$$mvr = \frac{nh}{2\pi}, \quad \text{т. е.} \quad mvr = n\hbar. \tag{4}$$

Это условие совпадает с боровским правилом квантования орбит электрона в атоме водорода. Каждому стационарному состоянию соответствует свое целочисленное значение *квантового числа* n.

Уровни энергии атома водорода. В планетарной модели атома водорода правило квантования (4) позволяет сразу найти значения энергии атома в стационарных квантовых состояниях. Используя второй закон Ньютона для электрона, обращающегося вокруг ядра под действием силы кулоновского притяжения (в гауссовой системе единиц)

$$\frac{mv^2}{r} = \frac{e^2}{r^2}, \tag{5}$$

и выражение для полной энергии электрона в атоме, состоящей из кинетической энергии орбитального движения и потенциальной энергии взаимодействия с ядром

$$E = \frac{mv^2}{2} - \frac{e^2}{r}, \tag{6}$$

получаем с помощью (4) радиусы r_n разрешенных орбит и энергии E_n стационарных состояний:

$$r_n = \frac{\hbar^2}{me^2} n^2 = a_0 n^2, \tag{7}$$

$$E_n = -\frac{me^4}{2\hbar^2} \frac{1}{n^2} = -R \frac{1}{n^2}. \tag{8}$$

Здесь $R = me^4/(2\hbar^2) = 13{,}6$ эВ — *энергия ионизации* атома водорода, т. е. минимальная энергия, которую нужно сообщить атому водорода в основном состоянии (в состоянии с наименьшей энергией с $n = 1$) для того, чтобы удалить из него электрон. Величина

$$a_0 = \frac{\hbar^2}{me^2} = 0{,}53 \cdot 10^{-8} \text{ см} \tag{9}$$

называется *боровским радиусом*. Это радиус ближайшей к ядру орбиты электрона в планетарной модели атома. Он характеризует размер атома водорода в основном состоянии с энергией

$$E_1 = -\frac{me^4}{2\hbar^2} = -R = -13{,}6 \text{ эВ}, \tag{10}$$

в котором свободный атом может находиться сколь угодно долго. Время жизни свободного атома в возбужденных состояниях составляет примерно 10^{-8} с.

Спектральные серии. Энергия квантов света, испускаемого возбужденными атомами водорода при переходах в более низкие энергетические состояния, определяется, в соответствии с формулой (1), разностью энергий начального и конечного состояний:

$$h\nu = E_n - E_m = R\left(\frac{1}{m^2} - \frac{1}{n^2}\right). \tag{11}$$

Переходы на уровень с $m=2$ из состояний $n=3, 4, \ldots$ образуют упоминавшуюся выше серию спектральных линий Бальмера, приходящуюся на видимую область спектра. Переходы на уровень с другим фиксированным значением m образуют другие серии спектральных линий. Например, переход на уровень с $m=1$ образуют серию Лаймана, лежащую в ультрафиолетовой области спектра. Во всех случаях наблюдается хорошее согласие между положением линий в спектре и предсказаниями по формуле (11).

Атом Бора и принцип соответствия. Полученные с помощью постулатов Бора формулы для частот спектральных линий, излучаемых атомом водорода, противоречат представлениям классической физики, согласно которым частота излучаемого света должна совпадать с частотой обращения электрона вокруг ядра. Однако они удовлетворяют принципу соответствия.

В применении к атому водорода принцип соответствия означает, что чем больше квантовое число n стационарного состояния, тем лучше выполняются для него законы классической физики. По мере увеличения n радиус орбиты электрона возрастает, а разность энергий двух соседних уровней стремится к нулю. При этом скачкообразные переходы между соседними уровнями становятся почти эквивалентными непрерывному процессу. В этом предельном случае результаты квантовой теории должны совпадать с результатами классической теории.

Применим формулу (11) для перехода атома между двумя соседними уровнями E_n и E_{n-1}, считая квантовое число n большим: $n \gg 1$. Тогда

$$\frac{1}{(n-1)^2} - \frac{1}{n^2} = \frac{n^2-(n-1)^2}{n^2(n-1)^2} \approx \frac{2}{n^3}$$

и (11) для частоты излучаемого света дает

$$\nu = \frac{me^4}{2\pi\hbar^3}\frac{1}{n^3}. \qquad (12)$$

Используя выражение (7) для радиуса орбиты электрона в n-м состоянии

$$r_n = n^2\frac{\hbar^2}{me^2},$$

перепишем формулу (12) в виде

$$\nu = \frac{e}{2\pi m^{1/2} r_n^{3/2}}. \qquad (13)$$

Но точно такое же выражение для частоты света, излучаемого электроном, обращающимся по круговой орбите радиуса r_n, дает классическая теория. Действительно, с точки зрения классической элект-

родинамики вращающийся по круговой орбите электрон должен излучать электромагнитные волны с частотой ν, равной частоте обращения электрона вокруг ядра. Применяя к движению электрона по круговой орбите с частотой ν второй закон Ньютона,

$$m \cdot 4\pi^2\nu^2 r_n = \frac{e^2}{r_n^2},$$

получаем

$$\nu = \frac{e}{2\pi m^{1/2} r_n^{3/2}}.$$

Итак, в области больших квантовых чисел квантовая теория дает тот же спектр излучения, что и классическая.

Современные представления о строении атома водорода. То, что на основе постулатов Бора удалось получить правильные количественные соотношения для атома водорода, связано с тем, что сами эти постулаты являются следствием фундаментальных положений современной квантовой теории.

Согласно квантовой механике бессмысленно говорить о движении электрона в атоме по определенной орбите. Физический смысл имеет только вероятность обнаружить электрон в том или ином месте. Квантовомеханическое распределение плотности вероятности местонахождения электрона в атоме можно представить в виде некоторого облака, окружающего ядро атома.

Каждое стационарное состояние характеризуется не одним квантовым числом n, называемым главным, а определенным набором квантовых чисел. Этому стационарному состоянию соответствует свое облако плотности вероятности, имеющее определенную пространственную конфигурацию. Одному и тому же значению главного квантового числа n, определяющего энергию атома $E_n = -me^4/(2\hbar^2 n^2)$, соответствует при $n > 1$ несколько различных состояний с одинаковой энергией, различающихся видом этого облака. Для состояний со сферически-симметричным облаком вероятность обнаружить электрон на некотором расстоянии от ядра имеет максимум, когда это расстояние равно радиусу соответствующей боровской орбиты

$$r_n = \frac{\hbar^2}{me^2} n^2.$$

Формулы для энергий стационарных состояний E_n и радиуса электронного облака r_n получены при использовании нерелятивистской механики, условие применимости которой состоит в малости скорости электрона по сравнению со скоростью света. Наибольшей скоростью

обладает электрон, находящийся в состоянии с $n = 1$, т. е. на языке теории Бора движущийся по ближайшей к ядру разрешенной орбите. Выразим скорость электрона v через радиус этой орбиты $a_0 = \hbar^2/(me^2)$. Уравнение движения по круговой орбите радиуса a_0 под действием кулоновской силы притяжения к ядру имеет вид

$$\frac{mv^2}{a_0} = \frac{e^2}{a_0^2},$$

откуда, подставляя a_0, находим

$$v^2 = \frac{e^2}{ma_0} = \left(\frac{e^2}{\hbar}\right)^2. \qquad (14)$$

С помощью (14) находим отношение v/c:

$$\frac{v}{c} = \frac{e^2}{\hbar c} = \alpha. \qquad (15)$$

Подставляя сюда числовые значения e, $\hbar$ и c, для безразмерной постоянной α получаем значение

$$\alpha = 7{,}3 \cdot 10^{-3} \approx \frac{1}{137}. \qquad (16)$$

Видно, что наибольшее возможное значение скорости электрона в атоме водорода в 137 раз меньше скорости света. Таким образом, этот атом представляет собой нерелятивистскую систему со сравнительно медленно движущимся электроном.

Постоянная тонкой структуры. Константа α играет фундаментальную роль в атомной физике. Она известна под названием *постоянной тонкой структуры*. Такое название объясняется тем, что впервые она появилась в физической теории при нахождении релятивистских поправок к уровням энергии в атоме, которые оказались пропорциональными $(v/c)^2 = \alpha^2$.

Постоянная тонкой структуры является одной из истинно фундаментальных констант природы, которая определяет не только релятивистские поправки, но и саму структуру атома. Поясним это. Сравним энергию кулоновского взаимодействия электрона с ядром при расстоянии между ними, равном боровскому радиусу a_0, с энергией покоя электрона mc^2:

$$\frac{e^2}{a_0 mc^2} = \frac{e^4}{\hbar^2 c^2} = \alpha^2.$$

Мы видим, что характерная для атома энергия выражается через энергию покоя электрона и постоянную тонкой структуры. Так как $\alpha^2 \ll 1$, то энергия связи электрона, т. е. энергия, необходимая для того, чтобы оторвать электрон от атома, много меньше энергии по-

коя электрона. Это означает, что атом представляет собой сравнительно «рыхлую», слабо связанную систему.

Взаимодействие между заряженными частицами, пропорциональное e^2, осуществляется через электромагнитное поле. Поэтому постоянная тонкой структуры $\alpha = e^2/\hbar c$ представляет собой естественный *безразмерный* параметр, характеризующий *интенсивность электромагнитного взаимодействия*. Ниже мы увидим, что именно этим параметром определяются особенности излучения электромагнитных волн (света) атомами.

Атомы с несколькими электронами. Состояния электронов в более сложных атомах, содержащих несколько электронов, характеризуется таким же набором квантовых чисел, что и у атома водорода. Однако здесь энергия электрона зависит не только от главного квантового числа n, характеризующего размер электронного облака в данном состоянии, но и от других квантовых чисел, характеризующих конфигурацию этого облака.

Атом с несколькими электронами представляет собой сложную систему взаимодействующих друг с другом электронов, движущихся в кулоновском поле ядра. Согласно квантовой механике в отсутствие внешних воздействий такая система может находиться только в определенных стационарных состояниях, характеризуемых дискретными значениями энергии. Строго говоря, при этом можно рассматривать только состояния всей системы в целом. Но тем не менее оказывается, что для атома можно приближенно говорить о состояниях каждого электрона в некотором эффективном поле со сферической симметрией, которое создается ядром и всеми остальными электронами. Ясно, что поле, действующее на некоторый электрон, зависит от состояния всех остальных электронов. Поэтому состояния всех электронов в атоме должны определяться одновременно.

Задача определения стационарных состояний электронов в атоме с несколькими электронами не может быть решена точно. Однако существуют различные приближенные методы, которые позволяют с очень высокой точностью рассчитывать наблюдаемые свойства атомов. Единственный атом, для которого эта задача может быть решена точно, — это атом водорода (или водородоподобный ион, т. е. ион с одним электроном, как, например, He^+, Li^{2+} и т. д.). Рассчитанные по квантовой механике уровни энергии атома водорода оказались совпадающими с теми, которые давала старая теория Бора. Успех теории Бора связан с тем, что она ввела в физику атома постоянную Планка в виде соотношения

$$rp \sim h,$$

связывающего положение и импульс электрона. Все выводы из теории Бора являются следствием этого соотношения, представляющего

собой частный случай соотношений неопределенностей Гейзенберга. Многие выводы, по существу, не связаны с используемой в теории Бора классической картиной движения электрона вокруг ядра.

Чтобы получить представление обо всей электронной оболочке атома, нужно рассмотреть, каким образом происходит заполнение электронами разрешенных стационарных состояний отдельных электронов. В атоме, находящемся в невозбужденном состоянии, электронные состояния должны быть заполнены таким образом, чтобы энергия всей его электронной оболочки имела наименьшее возможное значение. Конечно, наименьшее значение энергии получилось бы, если бы все электроны находились в наинизшем возможном состоянии, т. е. в состоянии с $n = 1$. Но опытные данные свидетельствуют о том, что заполнение электронной оболочки атома происходит иначе.

Принцип Паули. В 1925 г. В. Паули установил общий квантовомеханический принцип, согласно которому в любой системе в каждом разрешенном состоянии не может находиться более одного электрона.

Принцип Паули не имеет аналога в классической физике. Он связан с неразличимостью тождественных частиц в микромире: перестановка местами двух электронов в системе не может изменить ее состояния. Принцип Паули дает ключ к пониманию электронной структуры химических элементов и позволяет объяснить тот факт, что химические свойства некоторых атомов с разным числом электронов оказываются сходными. Выражением этого факта является периодичность химических свойств элементов, нашедшая отражение в эмпирически установленной периодической системе Д. И. Менделеева.

Многоэлектронные атомы. Попытаемся получить общее представление о строении тяжелых атомов, т. е. атомов с большим значением заряда ядра. Электронное облако таких атомов имеет слоистую, или, как говорят, оболочечную, структуру. На первый взгляд могло бы показаться, что размер электронного облака должен монотонно возрастать по мере перехода к все более и более тяжелым атомам, так что тяжелые атомы должны быть гораздо крупнее атома водорода. Однако в действительности это не так. Начнем с «голого» ядра и будем по одному последовательно прибавлять к нему все новые электроны. Для первого электрона, помещенного в поле ядра с зарядом Ze, можно воспользоваться теми же формулами, что и для атома водорода, заменив в них e^2 на Ze^2. Энергия связи такого электрона будет в Z^2 раз больше, чем в атоме водорода, а расстояние от ядра — в Z раз меньше радиуса a_0 первой боровской орбиты в атоме водорода. Это же будет приближенно справедливо и для второго прибавляемого электрона. Но что будет после прибавления n электронов?

На больших расстояниях от ядра поле такого иона будет похоже на поле ядра с зарядом $(Z - n)e$, так как ранее прибавленные электроны, находящиеся близко от ядра, частично экранируют его заряд. Поэтому легко понять, что последующие электроны оказываются все менее и менее сильно связанными с ядром. После прибавления предпоследнего, т. е. $(Z - 1)$-го, электрона электрическое поле иона будет подобно полю облака с суммарным зарядом $+e$. Радиус этого облака будет одного порядка с боровским радиусом a_0. Поэтому энергия связи последнего электрона будет сравнима с энергией связи электрона в атоме водорода, т. е. порядка десяти электрон-вольт, а окончательный размер многоэлектронного атома будет близок к размеру атома водорода и составит несколько ангстрем.

Даже такое грубое представление об электронной структуре тяжелого атома позволяет оценить, к какой области спектра относится излучение, связанное с переходами тех или иных электронов. Легко убедиться, что изменение состояния внешних электронов связано с излучением или поглощением электромагнитных волн оптического диапазона. Поэтому внешние электроны часто называют оптическими. Наоборот, переходы внутренних, близких к ядру атома электронов соответствуют ультрафиолетовой или рентгеновской области спектра.

Подводя итоги, мы приходим к выводу, что к находящимся в атомах электронам законы классической физики, вообще говоря, неприменимы. Такие фундаментальные свойства атомов, как их устойчивость, или особенности химического поведения, характеризуемые периодической системой элементов Д. И. Менделеева, противоречат классической физике и требуют для своего объяснения новых понятий современной квантовой физики.

- Покажите, что в предложенной Томсоном модели атома водорода на электрон, находящийся внутри атома, действует сила, пропорциональная его смещению из центра атома.
- Почему в опытах по рассеянию альфа-частиц атомами электроны, входящие в состав атомов, не играют никакой роли?
- Покажите, что положительный заряд, распределенный по всему объему атома, не мог бы отклонить альфа-частицу в опытах Резерфорда на большие углы. Оцените наибольший возможный угол рассеяния альфа-частицы в такой модели распределения заряда в атоме.
- Почему ядерная модель атома может быть только динамической? Другими словами, почему в рамках такой модели электроны в атоме не могут покоиться?
- Почему в планетарной модели атома гравитационная сила притяжения электронов к ядру никакой роли не играет?
- Объясните, в чем постулаты Бора противоречат классической механике и электродинамике.

- Поясните связь правила квантования с гипотезой де Бройля о волновых свойствах электрона.
- Получите формулы (7) и (8), определяющие радиусы орбит и энергии в стационарных состояниях атома водорода.
- Что такое постоянная тонкой структуры? Какую роль она играет для объяснения строения и наблюдаемых свойств атома?
- Покажите, что формулы теории Бора для частот спектральных линий атома водорода удовлетворяют принципу соответствия.
- Сформулируйте принцип Паули и объясните, каким образом он используется для объяснения периодичности химических свойств различных элементов.
- Почему размеры тяжелых многоэлектронных атомов практически такие же, что и у атома водорода?

§ 11. Измерения в атомной физике

Познакомившись с основными представлениями о строении атома, остановимся теперь на тех экспериментальных фактах, которые послужили фундаментом для развития этих представлений.

Опыты Дж. Томсона. Удельный заряд частицы, т. е. отношение электрического заряда частицы к ее массе, можно определить, изучая отклонение пучка движущихся частиц поперечными электрическим и магнитным полями. Измерение удельного заряда электрона таким методом было впервые выполнено Дж. Томсоном с помощью катодных лучей. Катодные лучи в действительности представляют собой пучок быстрых электронов в трубке, из которой выкачан воздух.

Если бы все частицы в пучке имели одинаковую скорость, то определить отношение их заряда e к массе m можно было бы следующим способом. Отклонение z пучка частиц, параллельного пластинам конденсатора, электрическим полем напряженности E (рис. 25), определяется соотношением

$$z = \frac{at^2}{2} = \frac{1}{2}\frac{eE}{m}\left(\frac{l}{v}\right)^2. \tag{1}$$

Измеряя z, можно было бы найти удельный заряд e/m, если бы была известна скорость частиц v. Эту скорость можно определить, если, кроме электрического создать магнитное поле, перпендикулярное пучку частиц и электрическому полю, и подобрать его индукцию B так, чтобы пучок не отклонялся. Компенсация действия электрической и магнитной сил будет происходить при выполнении условия

$$eE = evB, \tag{2}$$

откуда $v = E/B$.

Именно таким путем Дж. Томсон впервые измерил e/m для катодных лучей в 1897 г. Оказалось, что для катодных лучей при любом составе остаточного газа в трубке и при любом материале электродов

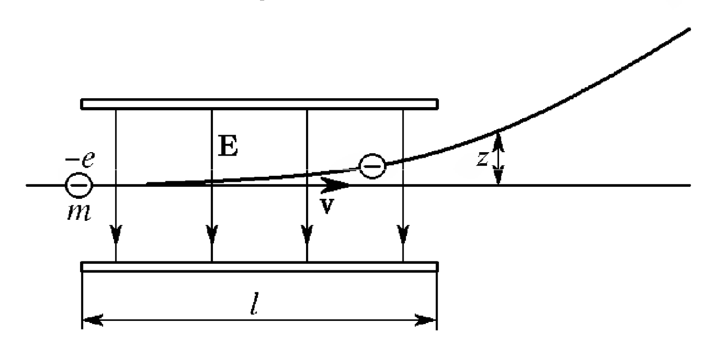


Рис. 25. К расчету отклонения пучка электронов в однородном электрическом поле

отношение e/m имеет одно и то же значение. Установление этого факта и было по существу экспериментальным открытием электрона.

Опыты Х. Буша. Одна из трудностей в этих исследованиях состояла в том, что электроны в пучке катодных лучей обладали большим разбросом скоростей. Для исключения влияния этого разброса Томсон в других модификациях опыта использовал параллельные электрическое и магнитное поля. После прохождения через такие поля частицы с единственным значением e/m и различными скоростями образуют на фотопластинке след в виде параболы. По форме этой параболы можно было определить e/m со значительно большей точностью. Применяя такой «метод парабол» к ионным пучкам (анодным лучам), Дж. Томсон в 1912 г. открыл изотопы нерадиоактивных элементов.

Определение удельного заряда электрона можно выполнить с использованием только магнитного поля. Такие измерения были впервые выполнены в 20-е гг. Х. Бушем методом фокусировки пучка электронов в продольном магнитном поле. Продольная фокусировка расходящегося пучка электронов была рассмотрена в § 16 книги 2 курса. Этот метод позволяет выполнить измерения с большей точностью, но предъявляет более жесткие требования к характеристикам электронных пучков.

Опыты Р. Милликена. Из опытов с ионными пучками можно определить не только e/m, но и сам электрический заряд, если из независимых источников получить информацию о массе иона. Однако такими методами не удается получить точного значения элементарного электрического заряда e.

Точные измерения элементарного заряда были выполнены Р. Милликеном в серии классических опытов 1908—1916 гг.

Эти опыты принесли неопровержимое доказательство дискретности электрических зарядов в природе.

Милликен измерял электрический заряд малых капелек масла, используемого для смазки часовых механизмов. Схема его установки показана на рис. 26. В тщательно изготовленный плоский конденсатор через отверстие в верхней пластине могут попадать мелкие капли масла, полученные с помощью специального распылителя. Чтобы конвекционные потоки воздуха не влияли на поведение капель, конденсатор помещался в защитный кожух, температура и давление внутри которого поддерживались постоянными.

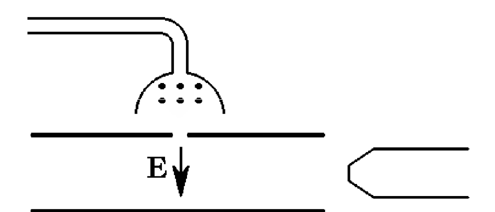


Рис. 26. Схема опытов Милликена

На пластины конденсатора подавалось постоянное напряжение от источника в несколько киловольт. В ходе опыта это напряжение можно было изменять. При распылении масла образующиеся капли, как правило, заряжаются и, попадая в конденсатор, движутся под действием силы тяжести и приложенного электрического поля. Движение отдельной капли можно наблюдать в микроскоп через специальное окошко.

С помощью рентгеновского излучения можно ионизировать воздух между пластинами конденсатора. Тогда заряд капли, а с ним и скорость установившегося движения в воздухе в том же электрическом поле могут скачком измениться. Измеряя скорости устоявшегося движения одной и той же капли в одном и том же электрическом поле, можно сравнить ее прежний и новый заряды q и q'. Если капля мала, а электричество имеет атомистическое, дискретное строение, то можно ожидать, что заряд капли состоит из небольшого числа элементарных зарядов. В таком случае отношение q'/q будет отношением небольших целых чисел. Именно это и наблюдалось в опытах Милликена.

Дискретность электрического заряда. В результате было надежно установлено, что в природе любые электрические заряды состоят из дискретных порций определенной величины. Такая величина обозначается через e и называется *элементарным электрическим зарядом*. Его числовое значение равно $1{,}602 \cdot 10^{-19}$ Кл. Заряд электрона равен $-e$. Точно таким же положительным зарядом обладает позитрон. Еще более замечательным фактом является равенство по величине заряда электрона и положительного заряда протона. Последнее равенство было проверено в очень тонких экспериментах, которые заключались в исследовании электронейтральности атома водорода и молекулы водорода. Их результаты свидетельствуют, что электрон и протон имеют одинаковые заряды с погрешностью до 10^{-20}.

По современным представлениям электрон и протон — это совершенно разные, очень далекие друг от друга по своим характеристикам элементарные частицы. Совпадение их электрических зарядов с такой фантастической степенью точности не имеет в современной физике

никакого глубокого теоретического обоснования. Квантование электрического заряда и по сей день представляется таинственным универсальным законом природы. Никто не знает, почему электрические заряды всех заряженных элементарных частиц в точности одинаковы.

Опыт показывает также, что электрические заряды всех заряженных элементарных частиц не только одинаковы, но и не зависят от состояния их движения. Электрическая нейтральность установлена на опыте и для атома гелия с почти такой же высокой точностью, как и для молекулы водорода. Атом гелия построен из тех же заряженных «кирпичиков», что и молекула водорода, — двух протонов и двух электронов, — но движутся эти частицы в атоме гелия совершенно иначе, нежели в молекуле водорода.

«Взвешивание» электрона. Опыты Милликена, в которых был измерен элементарный электрический заряд, можно рассматривать и как «взвешивание» электрона, массу которого непосредственно определить невозможно. Измерив заряд электрона, Милликен смог определить его массу, воспользовавшись найденным Томсоном отношением заряда к массе в опытах по отклонению электронных пучков. Ситуация здесь до некоторой степени аналогична «взвешиванию» Земли в опыте Кавендиша по лабораторному измерению гравитационной постоянной: взвесить Землю непосредственно невозможно, но можно определить произведение ее массы на гравитационную постоянную из астрономических наблюдений, т. е. по отклонению движущихся тел гравитационным полем Земли.

Массы атомов и молекул. Наиболее точные методы измерения масс атомов и молекул основаны на фокусировке пучков заряженных частиц — ионов этих атомов или молекул — при помощи раз-

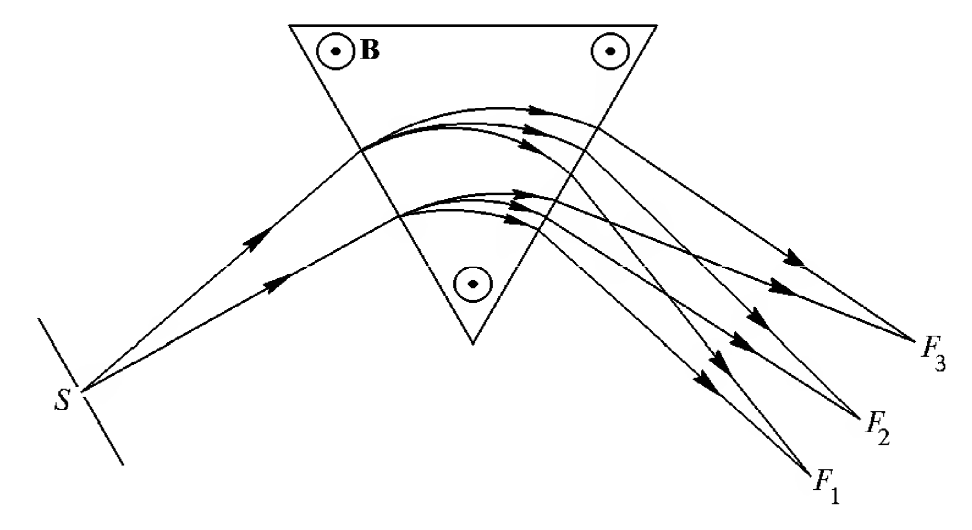


Рис. 27. Схема масс-спектрометра с фокусировкой однородным магнитным полем

личных комбинаций электрических и магнитных полей. Устроенные на таком принципе приборы называются масс-спектрографами (в

случае фотографической регистрации) и масс-спектрометрами (при электрической регистрации). Первый масс-спектрограф был изобретен и построен Ф. Астоном.

Существует большое разнообразие конфигураций полей, используемых в этих приборах. В наиболее распространенном масс-спектрографе с однородным магнитным полем ионы, образованные в ионном источнике, вылетают из щели S (рис. 27) в виде расходящегося пучка. Распространяясь в вакууме, этот пучок в магнитном поле разделяется на пучки ионов с разными значениями отношения заряда к массе. Каждый такой пучок, несмотря на разброс скоростей ионов, магнитное поле собирает в своем фокусе, т. е. фокусировка пучков частиц с разными значениями e/m происходит в разных местах F_1, F_2, F_3, ... Сравнивая положения фокусов, можно определить отношения масс ионов в разных пучках.

Принцип действия масс-спектрографа. Поясним принцип фокусировки расходящегося пучка заряженных частиц на примере метода фокусировки поперечным магнитным полем на протяжении половины окружности (π радиан).

Пусть источник ионов находится в точке S (рис. 28) и пусть все частицы характеризуются одним и тем же значением удельного заряда e/m. Будем считать, что все ионы вылетают из источника с одинаковой по модулю скоростью, но с небольшим разбросом по направлению. Ион, скорость которого направлена вверх на рис. 28, движется в магнитном поле, перпендикулярном плоскости рисунка, по окружности радиусом

$$R = \frac{mv}{eB}. \qquad (3)$$

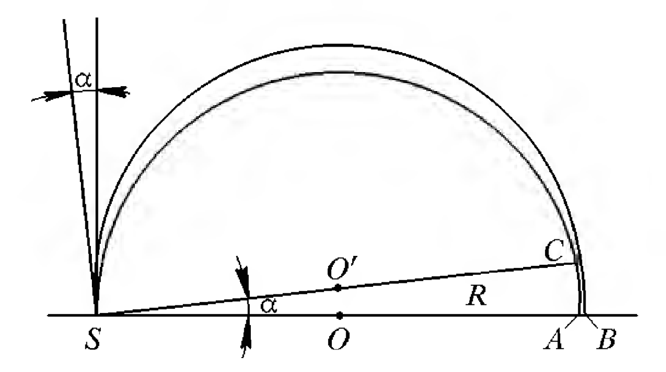


Рис. 28. Фокусировка поперечным однородным магнитным полем

Другой ион, вылетающий под небольшим углом α (рис. 28), в однородном магнитном поле движется по такой же окружности, повернутой как целое вокруг начальной точки S. Эта окружность пересечет диаметр SA первой окружности в точке B. Легко видеть, что

$$AB = 2R(1 - \cos\alpha) = 4R\sin^2\frac{\alpha}{2} \approx R\alpha^2. \qquad (4)$$

Хотя максимальное расхождение траекторий этих двух ионов, как видно из рис. 28, равно $R\alpha$, расстояние AB, пропорциональное квадрату малого угла α, значительно меньше. Поэтому, если источником S служит щель, через которую влетают ионы, то AB будет «изобра-

жением» этой щели в виде сравнительно узкой полоски. Таким образом, поперечное однородное магнитное поле фокусирует расходящийся пучок, действуя подобно цилиндрической линзе.

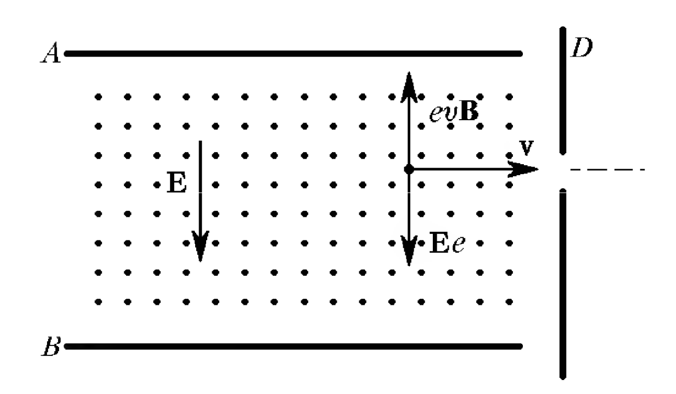


Рис. 29. Схема фильтра скоростей заряженных частиц

Селектор скоростей ионов. Для получения пучка ионов с одинаковыми по модулю скоростями используют фильтры скоростей. Схема самого простого фильтра скоростей показана на рис. 29: пучок заряженных частиц с различающимися скоростями пропускается через скрещенные электрическое и магнитное поля. Очевидно, что через малое выходное отверстие смогут пройти только те ионы, скорость которых удовлетворяет условию (2). Ионы с другими скоростями будут задержаны диафрагмой *D*.

Масс-спектрометр с поперечной фокусировкой. На рис. 30 показана схема масс-спектрометра А. Демпстера с фокусировкой на π радиан. Ионы, создаваемые источником *A*, приобретают одинаковую дополнительную энергию под действием ускоряющей разности потенциалов *U*, приложенной между источником *A* и выходной щелью *C*. Через

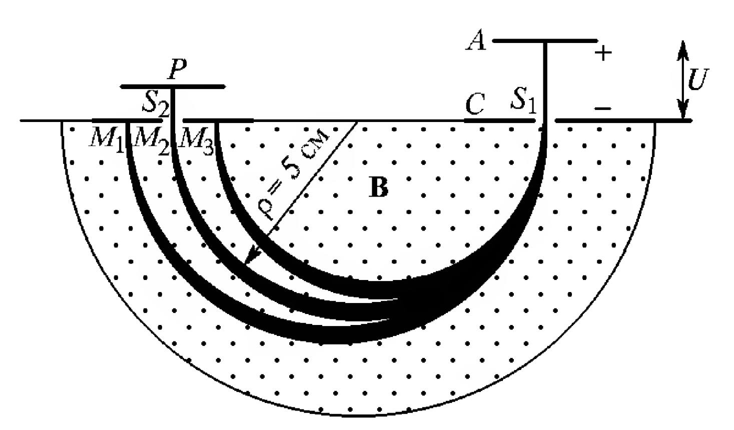


Рис. 30. Схема масс-спектрометра Демпстера

щель S_1 слегка расходящийся пучок ионов попадает в область, где действует поперечное однородное магнитное поле *B*. Ионы, обладающие различными значениями e/m, в соответствии с (3) фокусируются в разных местах, так как движутся по окружностям разных радиусов. Входящую в (3) скорость v можно опрсдслить из условия

$$\frac{mv^2}{2} = eU, \tag{5}$$

в котором пренебрегается первоначальной кинетической энергией теплового движения ионов по сравнению с энергией, приобретаемой ими при разгоне в электрическом поле. С помощью (3) и (5) для квадрата радиуса окружности получаем

$$R^2 = 2\frac{m}{e}\frac{U}{B^2}. \qquad (6)$$

Подбирая ускоряющее напряжение U или индукции магнитного поля B так, чтобы при данном m/e радиус R приобретал нужное значение, можно выпустить через выходную щель S_2 ионы любой массы. При этом можно не только измерять массы ионов, но и количественно исследовать состав исходного вещества: сила тока ионов через коллектор P служит мерой количества ионов данной массы.

Разделение изотопов. На таком же принципе действуют лабораторные и промышленные установки по разделению изотопов электромагнитным методом. Схема такой разделительной установки приведена на рис. 31. Установка размещается внутри большой вакуум-

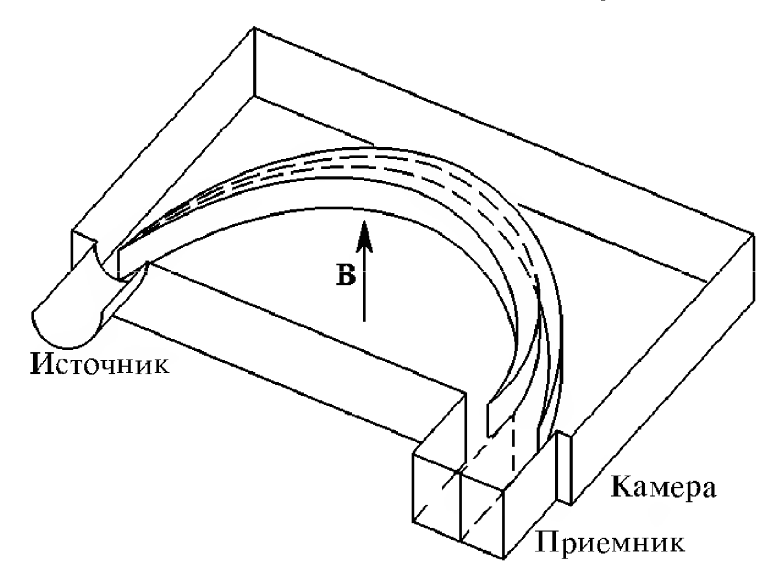


Рис. 31. Схема электромагнитной разделительной установки

ной камеры, помещенной между полюсами электромагнита. В отличие от измерительных масс-спектрометрических приборов в установках для разделения изотопов сила ионного тока в пучке в миллионы раз больше.

Электронная оптика. Движение пучков заряженных частиц в отклоняющих и фокусирующих системах масс-спектрографов можно рассматривать аналогично прохождению световых пучков в оптических приборах. На этой аналогии основана прикладная наука — электронная и ионная геометрическая оптика. Она получила большое развитие благодаря разнообразным практическим потребностям

в таких областях, как электронная микроскопия, телевидение, производство электронных микросхем и т. д.

В современных приборах, где используются электронные и ионные пучки, удается добиться совершенной фокусировки пучка частиц с заданным значением e/m, обладающих некоторым разбросом как модуля скорости, так и ее направления.

Разрешающая способность. Основной характеристикой качества масс-спектроскопической системы является ее разрешающая способность. Это понятие перенесено из теории оптических спектральных приборов. Говорят, что масс-спектрограф обладает разрешающей способностью r, если минимальная относительная разность масс $\Delta m/m$ для двух близких по массе частиц, регистрируемых раздельно, определяется равенством

$$\Delta m/m = 1/r.$$

В хороших современных масс-спектрометрах разрешающая способность достигает значения $8 \cdot 10^4$. На рис. 32 показана типичная спек-

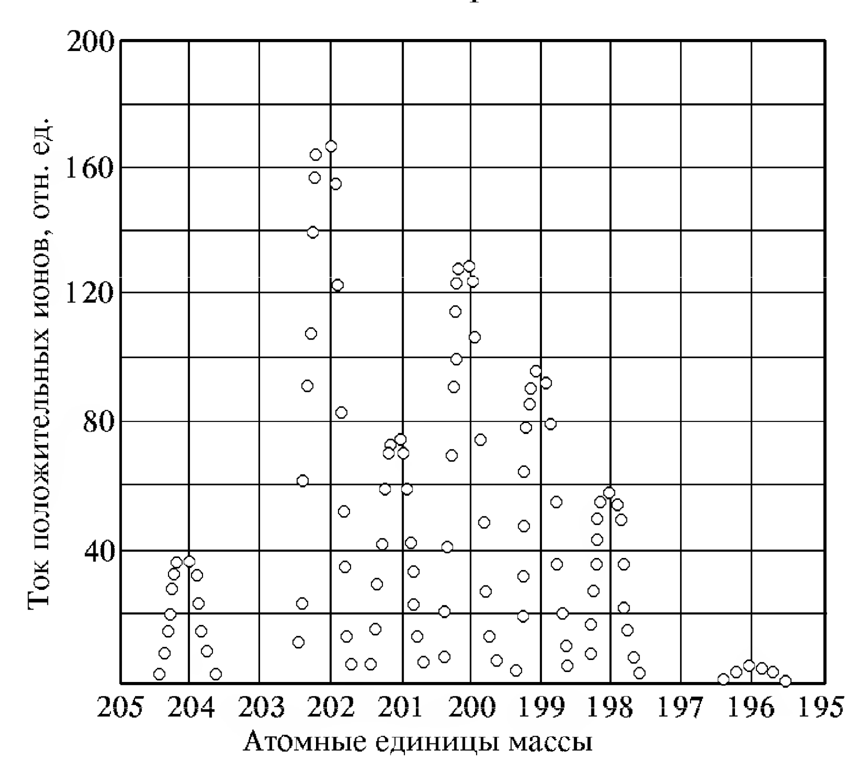


Рис. 32. Спектр масс изотопов ртути

трограмма, полученная на масс-спектрометре, предназначенном для изотопического анализа.

Атомная единица массы. Масс-спектрометрические методы лежат в основе наиболее точных измерений масс атомов и молекул. Так как в макроскопических масштабах массы атомов и молекул очень малы, в измерениях и расчетах удобно использовать не абсолютные

значения масс (в килограммах), а относительные. По международному соглашению массы всех атомов и молекул принято сравнивать с массой атома углерода, точнее его изотопа $^{12}_{6}C$. Углерод выбран потому, что он входит в огромное число химических соединений.

За атомную единицу массы принимается 1/12 массы изотопа углерода $^{12}_{6}C$, чтобы в этой шкале массы протона и нейтрона были близки к единице, а масса атома любого химического элемента выражалось числом, близким к целому.

Относительной атомной (или молекулярной) массой M_r вещества называют отношение массы атома (или молекулы) m_0 данного вещества к 1/12 массы $m_{0\,C}$ атома углерода $^{12}_{6}C$:

$$M_r = \frac{m_0}{(1/12)m_{0\,C}}. \qquad (7)$$

Относительные массы всех химических элементов измерены очень точно. Складывая относительные атомные массы элементов, входящих в состав молекулы определенного вещества, можно, зная его химическую формулу, приближенно вычислить относительную молекулярную массу. Например, относительная масса молекулы воды H_2O равна $2\cdot 1 + 16 = 18$, если эта молекула содержит наиболее распространенные изотопы $^{1}_{1}H$ и $^{16}_{8}O$.

Точные измерения атомных масс. Для точных определений неизвестной массы какого-либо иона необходимо построить градуировочную кривую масс-спектрометра. Эта кривая сначала строится грубо путем выбора цепочки масс атомных или молекулярных ионов, отличающихся приблизительно на одну единицу массы, например C, CH, CH_2, CH_3, CH_4. Если для данной цепочки массы ионов известны с погрешностью, например, 0,1% (что легко достигается химическими методами), то с такой же погрешностью можно построить график зависимости координаты линии в спектре масс на фотопластинке от относительной массы. Затем эту грубую градуировку уточняют. Это можно сделать, например, таким способом. Уменьшим ускоряющее напряжение при неизменном магнитном поле настолько, чтобы на место линии CH_3 попала линия CH_4. Тогда все новые положения массовых линий будут соответствовать значениям масс, изменившимся в отношении 16/15.

После такой градуировки можно измерить очень маленькую разность масс между некоторой неизвестной массой и эталонной. Среди множества углеводородов почти всегда можно подобрать очень близкие по массе соединения. Две близкие массы образуют так называемый дублет в спектре масс. Примером могут служить ионы $^{16}_{8}O$ и $^{12}_{6}C^{1}_{1}H_4$ (кислород и метан). Если относительная погрешность измере-

ния разности масс Δm этих ионов составляет всего 0,1% (что составляет несколько сотых долей от атомной единицы массы), то относительная погрешность измерения массы иона $^{16}_{8}O$ достигает одной миллионной. Иллюстрацией такого метода может служить рис. 33, на котором приведены два снимка нескольких близких массовых линий.

Обратим внимание на то, что этот дублет образован атомными числами, построенными из одинакового числа протонов, нейтронов и электронов. Различие в массах компонентов дублета обусловлено тем,

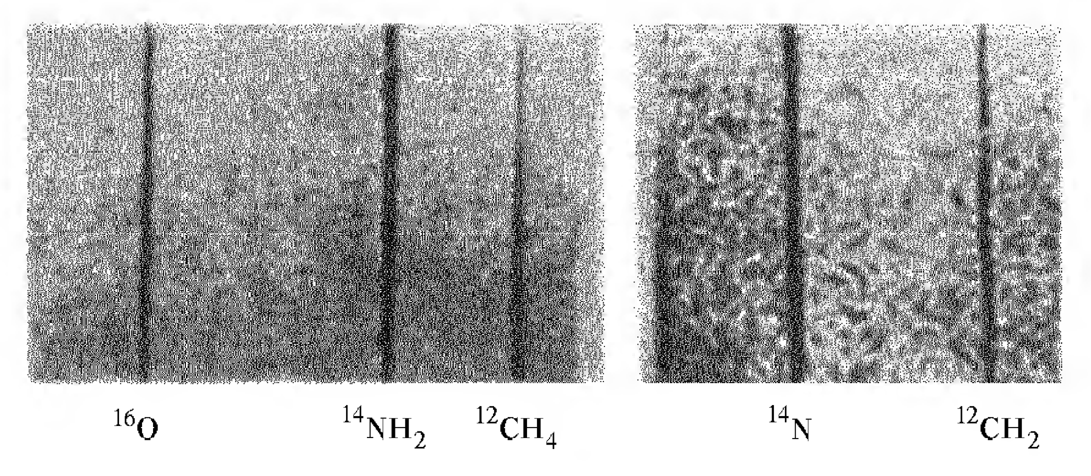


Рис. 33. Фотоснимки массовых линий

что у атома кислорода все протоны и нейтроны находятся в одном ядре, а в молекуле метана — в разных. Поэтому энергия взаимодействия протонов и нейтронов в этих системах различна, что приводит к различиям в энергии покоя и массе $^{16}_{8}O$ и $^{12}_{6}C^{1}_{1}H_4$. Масс-спектрометрические методы позволяют определять не только относительные, но и абсолютные значения масс атомов и молекул, хотя и со значительно меньшей точностью. Уменьшение точности обусловлено тем, что абсолютные измерения масс атомов выполняются косвенно, и на их вычисляемом значении сказываются как погрешности измерения параметров экспериментальной установки (геометрических размеров, напряженностей полей, ускоряющего напряжения и т. п.), так и точность, с которой известен элементарный электрический заряд e. Для массы атома углерода измерения дают значение $m_{0\,C} = 1{,}995 \cdot 10^{-26}$ кг.

Количество вещества. Постоянная Авогадро. Наиболее естественной единицей для измерения количества вещества в каждом теле могло бы служить число молекул или атомов. Однако число молекул в любом макроскопическом теле очень велико. Поэтому для характеристики количества вещества удобно использовать не абсолютное число молекул, а относительное.

В международной системе единиц (СИ) количество вещества измеряют в *молях*. Моль — это одна из основных единиц СИ: 1 моль —

это количество вещества, в котором содержится столько же молекул или атомов, сколько атомов содержится в углероде массой 0,012 кг.

Моль любого вещества содержит одно и то же число атомов или молекул. Это число обозначается символом N_A и называется *постоянной Авогадро*. Постоянную Авогадро можно найти, если разделить массу одного моля углерода на массу одного атома углерода:

$$N_A = \frac{0{,}012\ \text{кг/моль}}{1{,}995 \cdot 10^{-26}\ \text{кг}} = 6{,}02 \cdot 10^{23}\ \frac{1}{\text{моль}}. \tag{8}$$

Количество вещества ν, т. е. число молей, равно отношению числа молекул N в данном теле к постоянной Авогадро N_A:

$$\nu = \frac{N}{N_A}. \tag{9}$$

Молярная масса. Для характеристики конкретного вещества наряду с относительной молекулярной массой M_r в физике и химии широко используется *молярная масса* M. Молярной массой называют массу данного вещества, взятого в количестве одного моля. Молярная масса равна произведению массы m_0 одной молекулы данного вещества на постоянную Авогадро N_A:

$$M = m_0 N_A. \tag{10}$$

Если в формуле (9) числитель и знаменатель домножить на массу одной молекулы m_0, то для количества вещества ν в теле можно написать

$$\nu = \frac{m}{M}, \tag{11}$$

где $m = m_0 N$ — масса тела.

Размеры атомов и молекул. Представление о размерах отдельных атомов и молекул можно получить разными способами. Если предположить, что в конденсированном состоянии (т. е. в жидкой или твердой фазе) атомы или молекулы вещества «соприкасаются» друг с другом, то, определив приходящийся на одну молекулу объем, можно оценить ее размеры. Например, масса $m = 1$ г воды занимает объем $V = 1\ \text{см}^3$ и содержит $N = N_A m/M$ молекул, где $M = 18$ г/моль. Поэтому приходящийся на одну молекулу объем составляет $3 \cdot 10^{-23}\ \text{см}^3$, а ее линейный размер $r \sim \sqrt[3]{V} \approx 3 \cdot 10^{-8}$ см.

Согласующуюся с этим значением оценку размеров молекул можно получить без использования постоянной Авогадро, если обратиться к простому опыту по растеканию капли масла на поверхности воды (опыт Ленгмюра). Предположив, что в масляном пятне молекулы располагаются в один слой, можно оценить толщину слоя, а тем самым и линейный размер молекул масла, разде-

лив объем растекшейся капли масла на площадь образовавшегося пятна.

Более надежную информацию о расстояниях между атомами в твердых телах и жидкостях дают опыты по дифракции рентгеновского излучения — рентгено-структурный анализ. Оказывается,

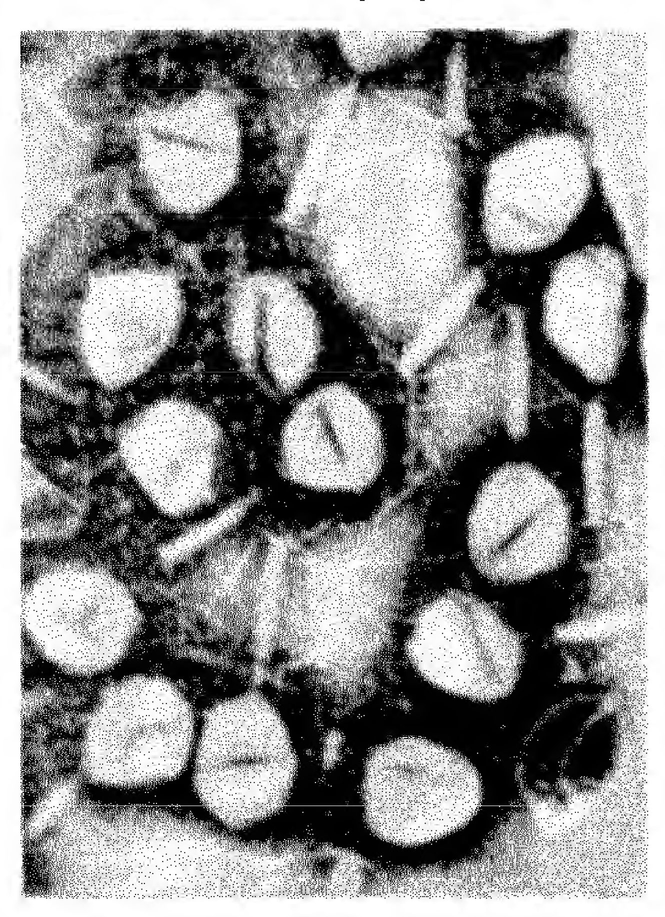


Рис. 34. Электронный снимок фага Т-2 с увеличением в 180 000 раз

что молекулы многих веществ имеют размеры порядка 10^{-8} см. Лишь молекулы полимеров и некоторых органических веществ (белков) могут иметь гигантские размеры, что позволяет получить их изображение в электронном микроскопе, на котором отчетливо различима их структура (рис. 34).

- Почему в опытах по отклонению пучка частиц в электрическом и магнитном полях можно определить не заряд частицы, а только отношение заряда к массе?
- Покажите, что след, оставляемый на фотопластинке пучком заряженных частиц, прошедших через поперечные, параллельные друг другу электрическое и магнитное поля, имеет форму параболы.
- Поясните идею метода фокусировки слабо расходящегося пучка электронов продольным магнитным полем.
- Милликен в своих опытах сначала использовал капельки воды, однако затем перешел к часовому маслу. Как вы думаете, с чем это связано?

- Почему установленная на опыте электрическая нейтральность молекулы водорода и атома гелия свидетельствует не только о равенстве электрических зарядов электрона и протона, но и о независимости величины заряда каждой частицы от ее движения?
- Проведите подробное сопоставление опыта Милликена по «взвешиванию» электрона и опыта Кавендиша по «взвешиванию» Земли.
- Поясните принцип фокусировки ионов в поперечном магнитном поле.
- Объясните принцип действия масс-спектрометра и его применение для количественного исследования состава вещества.
- Какую величину называют разрешающей способностью масс-спектрометра? Что она характеризует?
- Как вводится атомная единица массы? Почему она основывается на атоме углерода $^{12}_{6}C$?
- Что такое относительная атомная или молекулярная масса? Почему при ее вычислении с помощью химической формулы вещества мы получаем приближенные значения?
- Поясните, каким образом достигается относительная точность 10^{-6} при масс-спектрографическом измерении массы изотопа $^{16}_{8}O$ с использованием дублета $^{12}_{6}C^{1}_{1}H_4$ и $^{16}_{8}O$.
- При сгорании углерода 12 г углерода соединились с 32 г кислорода. Изменилось ли в результате этой химической реакции количество вещества?
- Сколько молекул воды содержится в капле объемом $1 см^3$?
- Как можно оценить размеры отдельных молекул?

§ 12. Молекулы

Все тела состоят из атомов или молекул. Что удерживает атомы в молекуле? Атомы соединяются в молекулы силами, которые обычно называют химическими. Конечно, никаких фундаментальных «химических» взаимодействий в природе не существует. Эти силы всегда сводятся к известным нам фундаментальным физическим взаимодействиям, причем определяющую роль в формировании химических связей играют электромагнитные силы. Проявления этих сил, несмотря на единую физическую природу, чрезвычайно разнообразны. Богатейший опыт изучения результатов действия сил между атомами накоплен в химии, но лишь с развитием квантовых представлений удалось дать объяснение разнообразным проявлениям этих сил на основе фундаментальных физических законов.

Химические силы и физика. Многие молекулы состоят всего из нескольких атомов. Однако существуют и гигантские молекулы, напри-

мер молекулы белков, которые содержат десятки или даже сотни тысяч атомов. Роль физики состоит в том, чтобы объяснить химическое взаимодействие атомов, т. е. то, как атомы соединяются в молекулы, а затем образуют макроскопические тела. Оказывается, что при переходе от микроскопических явлений к макроскопическим в некоторых случаях роль квантовых эффектов сохраняется, т. е. квантовая природа проявляется не только в микромире, но и в многочисленных свойствах макрообъектов. Например, высокая электропроводность меди и ее практически полное отсутствие у кварца объясняются в конечном счете именно квантовыми эффектами. Существование постоянных магнитов и сверхпроводников представляет собой особенно яркие примеры проявления квантовых эффектов в макромире.

Можно выделить два принципиально различных механизма, обеспечивающих связь атомов в молекулах: это так называемые *ионная* и *ковалентная* связи.

Ионная связь между атомами. Типичный представитель ионной связи — хорошо известная поваренная соль, называемая в химии хлористым натрием ($NaCl$). Каким образом два разнородных атома — атом металла натрия и атом газа хлора — образуют устойчивое соединение?

Атом натрия имеет один электрон на внешней оболочке; атом хлора имеет семь электронов на внешней оболочке, где в принципе может находиться восемь электронов, т. е. ему недостает одного электрона до заполнения внешней оболочки. Если один электрон атома натрия передать атому хлора, то и у того, и у другого образуются максимально устойчивые электронные конфигурации из целиком заполненных оболочек. Чтобы отнять электрон у атома натрия, нужно затратить энергию 5,1 электрон-вольта (эВ). При захвате этого электрона нейтральным атомом хлора образуется отрицательный ион Cl^-, и при этом освобождается энергия 3,7 эВ.

Казалось бы, образование молекулы Na^+Cl^- энергетически невыгодно, и такое соединение не может быть устойчивым. Однако между ионами с зарядами противоположных знаков существует электростатическое притяжение, и при их сближении потенциальная энергия взаимодействия убывает. При расстоянии между центрами двух ионов, равном 1,1 нм, убыль потенциальной энергии взаимодействия компенсирует затраты энергии в 1,4 эВ, необходимые для перехода электрона от атома натрия к атому хлора с образованием двух ионов. При дальнейшем сближении ионов появляется уже выигрыш в энергии, что приводит к более сильной связи ионов в молекуле.

Однако ионы могут сближаться лишь до некоторого предела (0,24 нм), соответствующего равновесному расстоянию между ионами в молекуле. На таком расстоянии кулоновские силы притяжения

уравновешиваются силами отталкивания, имеющими неэлектростатическое происхождение и связанными с проявлением принципа Паули. Дело в том, что при сближении ионов доступная для движения их электронов область пространства уменьшается, а это в свою очередь неизбежно ведет к увеличению кинетической энергии электронов. Это сугубо квантовый эффект, который можно объяснить с помощью соотношения неопределенностей Гейзенберга: при уменьшении области движения, т. е. неопределенности значения координат, неизбежно растет неопределенность значения импульса, а следовательно, и сам импульс.

Обусловленные принципом Паули силы отталкивания пренебрежимо малы, пока расстояние между ионами больше равновесного значения, и резко возрастают при попытке сблизить ионы на меньшее расстояние. Именно о таком характере зависимости сил отталкивания от расстояния говорит тот факт, что расчет энергии связи ионной молекулы NaCl, в котором учитывается только электростатическое взаимодействие ионов, дает хорошо согласующееся с опытом значение 5,5 эВ. Ситуация здесь аналогична распределению потенциальной энергии упругой деформации двух последовательно соединенных пружин разной жесткости: в основном энергию запасает менее жесткая пружина, т. е. та, у которой меньше коэффициент жесткости k. Это обстоятельство позволяет считать ионы твердыми шарами определенного радиуса.

Ковалентная связь. Ионная связь, возникающая в результате перехода одного или нескольких электронов от одного атома к другому, характерна для соединений, в которых участвуют химические элементы из противоположных концов периода системы Менделеева. Ковалентная связь возникает вследствие обобществления атомами одного или нескольких электронов и характерна для одинаковых или близких по своим химическим свойствам атомов.

В простейшем виде ковалентная связь реализуется в молекулярном ионе водорода H_2^+, в котором оба протона имеют всего один общий электрон. Электростатическое отталкивание протонов стремится развалить молекулу. Но, с другой стороны, энергия связи электрона в поле двух протонов, конечно, больше, чем при наличии только одного протона. При определенном расстоянии между протонами (0,1 нм) второй фактор оказывается более существенным, что обеспечивает устойчивость такой системы. Для удаления одного из протонов и образования атома водорода в основном состоянии требуется энергия 2,55 эВ. Еще прочнее оказывается ковалентная связь в нейтральной молекуле водорода (4,48 эВ), где она образуется двумя электронами.

Для ковалентной связи характерна определенная пространственная направленность и насыщаемость. У атома углерода есть четыре электрона, способные участвовать в образовании ковалентных свя-

зей. Простейший пример углеродных соединений дает молекула метана CH_4, в которой пространственная направленность ковалентных связей проявляется в том, что четыре атома водорода располагаются в вершинах правильного тетраэдра, в центре которого находится атом углерода, так что углы между всеми валентными связями одинаковы (рис. 35). Насыщенность связей проявляется в том, что больше четырех атомов водорода атом углерода присоединить не может.

В действительности в природе практически не встречается чисто ионных и чисто ковалентных связей, и можно говорить лишь о преимущественно ионном или ковалентном характере той или иной связи. Преимущественно ковалентная химическая связь характерна для таких молекул, как H_2, CO и т. д., а также для большинства органических молекул; ионная — для молекул Cs, KCl, многочисленных ионных кристаллов.

Двухатомные молекулы, состоящие из разных атомов, обладают электрическим дипольным моментом, так как у них центры положительного и отрицательного зарядов смещены относительно друг друга вдоль соединяющей ядра атомов прямой.

Взаимодействие атомов в молекуле. Рассмотрим подробнее взаимодействие образующих молекулу атомов. Несмотря на всю слож-

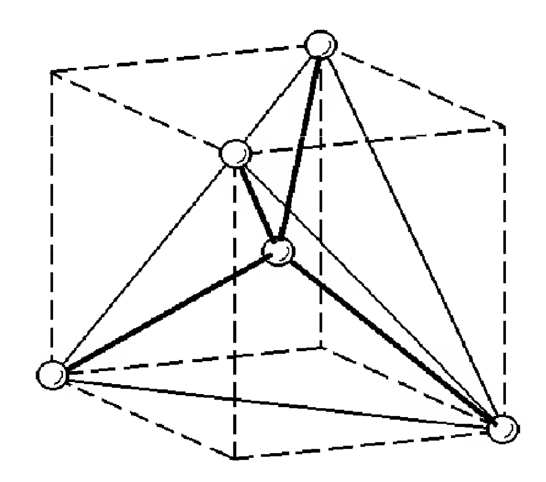


Рис. 35. Схематический рисунок, иллюстрирующий строение молекулы метана. Атомы водорода расположены в вершинах правильного тетраэдра, или в четырех из восьми вершин куба, в центре которого находится атом углерода

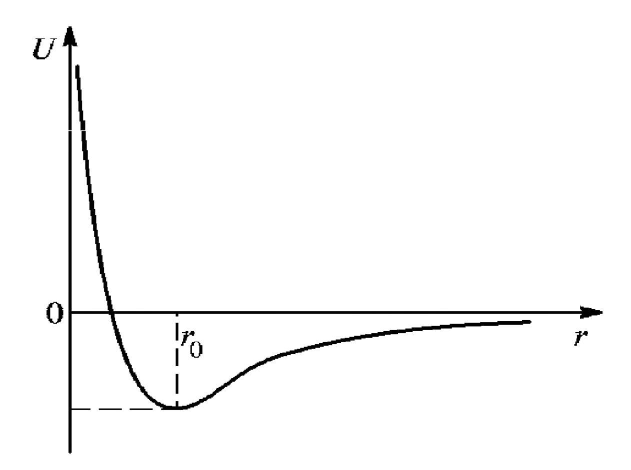


Рис. 36. Зависимость потенциальной энергии взаимодействия атомов в двухатомной молекуле от расстояния между ними

ность упомянутых выше сил притяжения и отталкивания, действующих между электронами и ядрами этих атомов, можно сравнительно просто феноменологически описать результирующее взаимодействие атомов в молекуле с помощью графика зависимости потенциальной энергии от расстояния между атомами.

При любой природе сил кривая потенциальной энергии двухатомной молекулы имеет характерный вид, показанный на рис. 36. Эта кривая описывает все детали взаимодействия между атомами. Для находящихся на больших расстояниях атомов преобладает сила притяжения, причем эта сила быстро уменьшается с увеличением расстояния между ними. При сближении атомов крутизна графика (наклон касательной) увеличивается, что говорит о возрастании силы притяжения атомов. При дальнейшем сближении атомов сила притяжения начинает убывать и обращается в нуль при некотором значении r_0 расстояния между ними. В этой точке потенциальная энергия имеет минимум. На расстояниях, меньших r_0, результирующая сила уже имеет характер отталкивания, причем она резко возрастает при сближении атомов, о чем свидетельствует быстрый рост крутизны графика $U(r)$ левее точки r_0.

Колебания атомов в молекуле. Расстояние r_0 соответствует устойчивому положению равновесия атомов. На таком расстоянии друг от друга в молекуле находятся неподвижные атомы. В непосредственной окрестности точки r_0 кривую $U(r)$ можно аппроксимировать параболой, т. е. считать зависимость потенциальной энергии от смещения из этого положения квадратичной. Это означает, что движение атомов в молекуле при малых смещениях из положения равновесия, рассматриваемое на основе классической механики, представляет собой почти гармоническое колебание. Но в отличие от потенциальной энергии гармонического осциллятора потенциальная яма на графике $U(r)$ в действительности асимметрична. Это значит, что с увеличением амплитуды колебаний они становятся ангармоническими. Ангармонизм колебаний проявляется в том, что среднее расстояние между атомами, а тем самым и средний размер молекулы, возрастает с увеличением энергии колебаний.

Глубина потенциальной ямы на рис. 36 характеризует энергию связи атомов в молекуле, т. е. энергию, необходимую для того, чтобы развести атомы и тем самым разрушить молекулу. Эта энергия определяет тепловой эффект химических реакций. При таких реакциях из молекул одного сорта образуются другие молекулы, т. е. одни химические связи разрываются, а другие возникают. Что больше — работа разрыва или работа образования новых связей? В природе мы сталкиваемся с реакциями обоих типов.

Энергия химических превращений. Тепловые эффекты реакций, т. е. результирующее выделение или поглощение энергии — это большей частью величины порядка нескольких сотен килоджоулей на моль, что составляет около одного электронвольта на одну молекулу.

Например, реакция сгорания углерода (графита) с учетом выделения энергии записывается так:

$$C + O_2 = CO_2 + 4{,}1\ \text{эВ}, \qquad (1)$$

т. е. при образовании каждой молекулы углекислого газа выделяется энергия 4,1 эВ.

С помощью таких уравнений можно найти тепловые эффекты химических превращений, для которых по тем или иным причинам не годятся прямые способы измерений. Например, если бы два атома углерода (графита) соединились с одной молекулой водорода, то образовался бы газ ацетилен:

$$2C + H_2 = C_2H_2 + ?\ . \qquad (2)$$

Напрямую такая реакция не идет. Тем не менее можно рассчитать ее тепловой эффект. Наряду с рассмотренной выше реакцией окисления углерода запишем еще две известные реакции окисления водорода и ацетилена:

$$2H_2 + O_2 = 2H_2O + 5{,}9\ \text{эВ}, \qquad (3)$$

$$2C_2H_2 + 5O_2 = 4CO_2 + 2H_2O + 27\ \text{эВ}. \qquad (4)$$

Эти равенства можно по существу трактовать как уравнения для энергии связи молекул и оперировать с ними, как с обычными алгебраическими равенствами. Вычтем почленно из (4) сначала равенство (3), а затем умноженное на 4 равенство (1). Тогда получим

$$2C + H_2 = C_2H_4 - 2{,}35\ \text{эВ}. \qquad (5)$$

Это означает, что для интересующего нас превращения двух атомов углерода и молекулы водорода в молекулу метана необходимо затратить энергию 2,35 эВ.

Взаимодействие между молекулами. Между молекулами действуют силы, очень похожие на силы между атомами, о которых говорилось выше. Эти силы также имеют электромагнитную природу. Кривая потенциальной энергии взаимодействия атомов (см. рис. 36) качественно правильно передает и основные черты взаимодействия молекул. Однако между этими взаимодействиями имеются и существенные различия.

Прежде всего, существенно различаются равновесные расстояния между атомами одной молекулы и атомами двух соседних взаимодействующих молекул. Например, атомы кислорода в молекуле O_2 находятся на расстоянии 0,12 нм. Атомы кислорода разных молекул в затвердевшем кислороде находятся друг от друга на расстоянии 0,29 нм.

Большое различие есть не только в равновесных расстояниях, но и в глубине потенциальной ямы: для «чужих» атомов эта яма существенно мельче. Поэтому молекулы легче оторвать одну от другой, чем разделить молекулу на атомы. Если энергия, необходимая для разрыва связей между атомами кислорода, образующими молекулу, составляет около 4,3 эВ, то для удаления на большое расстояние двух молекул кислорода необходимая энергия в 50 раз меньше.

Однако различия межатомных сил в молекуле и молекулярных сил проявляются не только в количественных характеристиках потенциальной кривой. Как уже отмечалось выше, атомы соединяются в молекулу с вполне определенным числом других атомов. Если два атома водорода образовали молекулу, то третий уже не может присоединиться к ним для этой цели. Ничего подобного мы не находим в межмолекулярном взаимодействии. Здесь отсутствует насыщаемость сил взаимодействия. Притянув к себе одного соседа, молекула ни в коей мере не утрачивает способности притягивать других.

Модельные потенциалы. Квантовая механика на основе известного строения атомов позволяет приближенно рассчитывать потенциальные кривые их взаимодействия в молекулах. Однако сделать это для межмолекулярного взаимодействия значительно сложнее. Поэтому для его теоретического описания обычно исходят не из первых принципов, а используют некоторые модельные потенциалы, подбирая входящие в выражение для $U(r)$ параметры таким образом, чтобы расчеты на основе этих потенциалов хорошо согласовывались с экспериментальными данными. При таком феноменологическом подходе чаще всего пользуются потенциалом Леннарда—Джонса:

$$U(r) = -\frac{a}{r^6} + \frac{b}{r^{12}} \qquad (a,\, b > 0). \tag{6}$$

Первое слагаемое описывает притяжение между молекулами, преобладающее на больших расстояниях r. Зависимость сил притяжения от расстояния такая же, как у двух электрических диполей, хотя эта модель может применяться и для молекул, не обладающих собственным постоянным дипольным моментом. Второе слагаемое в правой части (6), имеющее более сильную зависимость от r, описывает преобладающее на малых расстояниях отталкивание молекул. К хорошему согласию с данными экспериментов приводит допущение, что потенциальная энергия сил отталкивания возрастает с уменьшением расстояния именно по закону r^{-12}, что соответствует силе отталкивания, пропорциональной r^{-13}. Силы отталкивания связаны с принципом Паули и не имеют наглядного классического объяснения. Качественный вид потенциальной кривой, описываемой формулой (6), совпадает с показанным на рис. 36. Параметры a и b в (6) связаны простыми соотношениями с положением и глубиной

потенциальной ямы. Они определяются из различных экспериментальных данных, полученных при измерении диффузии, теплопроводности, вязкости и т. д.

Проведенные в последние годы квантовомеханические расчеты позволили обосновать используемые модельные потенциалы и в значительной степени разобраться в физической природе сил межмолекулярного взаимодействия.

Молекула в химии и в физике. Представления о том, что такое молекула, в физике и в химии несколько различаются. В химии под молекулами понимают наименьшие одинаковые структурные элементы вещества, из которых оно может быть построено. Например, с точки зрения химика поваренная соль состоит из молекул NaCl. Однако в действительности отдельно от макроскопического образца вещества такая молекула существовать не может. В кристалле поваренной соли у каждого иона Na^+ имеется 6 ближайших соседей — ионов Cl^-, находящихся от него на одинаковом расстоянии (рис. 37). Со всеми шестью соседями ион Na^+ связан совершенно одинаково. Так как ни одному из соседей нельзя отдать предпочтение, то в определенном смысле весь

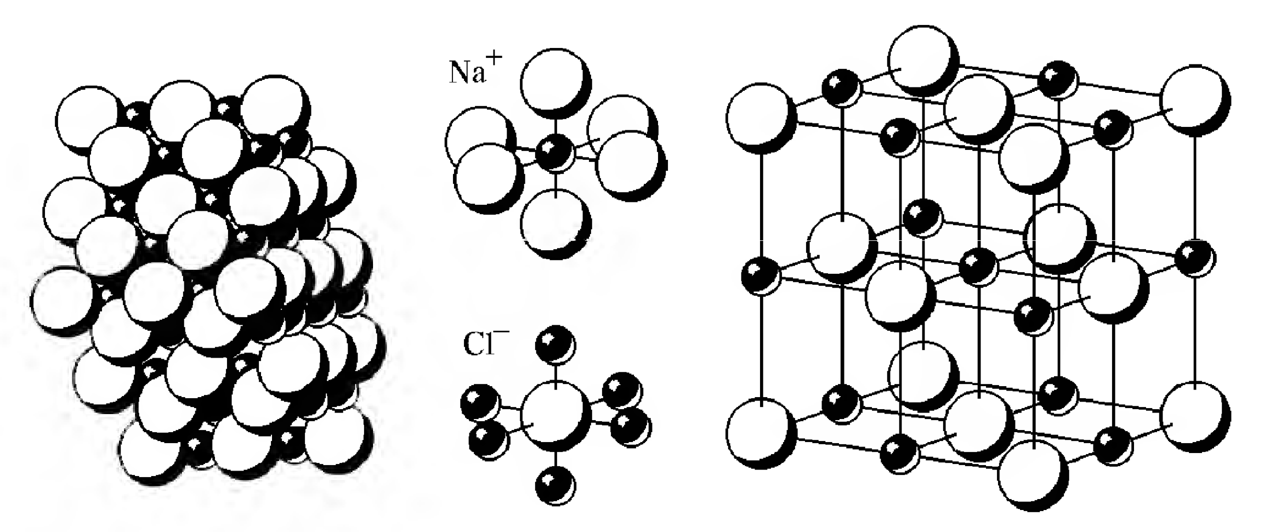


Рис. 37. В кубической решетке кристалла NaCl каждый ион натрия или хлора окружен шестью ионами противоположного знака

кристалл NaCl следует рассматривать как единую гигантскую молекулу (сверхмолекулу).

В физике под молекулами понимают наименьшие одинаковые структурные образования, сохраняющие свою индивидуальность при разделении вещества на такие мельчайшие части. При таких превращениях вещества, как растворение, плавление, испарение, физические молекулы не разрушаются. Например, в стакане сладкого чая «плавают» молекулы сахара. А вот в соленой воде никаких молекул поваренной соли не существует: все структурные связи NaCl в растворе разрушаются и поваренная соль в растворе существует в виде отдельных положительных ионов Na^+ и отрицательных ионов Cl^-.

- Существует ли какой-то особый вид химических сил, удерживающих атомы в молекулах, или и здесь все сводится к известным в физике фундаментальным взаимодействиям? Если да, то к каким именно?
- Покажите, что энергия, которую нужно затратить для того, чтобы развести на бесконечность ионы Na^+ и Cl^- с расстояния 1,1 нм, составляет 1,4 эВ?
- Поясните аналогию сил притяжения и отталкивания между ионами в молекуле с силами упругости последовательно соединенных пружин разной жесткости. Покажите, что запасаемая в них потенциальная энергия обратно пропорциональна жесткости.
- Чем объясняется пространственная направленность и насыщаемость ковалентных связей?
- По графику зависимости потенциальной энергии двухатомной молекулы от расстояния (рис. 36) постройте качественный график зависимости от расстояния действующей между атомами силы.
- Как выбрано начало отсчета потенциальной энергии на графике $U(r)$, показанном на рис. 36?
- Поясните, почему движение атомов в молекуле при малых смещениях представляет собой гармоническое колебание. Чем определяется частота этого колебания? Как она связана с крутизной кривой $U(r)$ вблизи дна потенциальной ямы?
- Почему при учете ангармонизма колебаний атомов средний размер молекулы увеличивается с ростом энергии колебаний? Где при колебаниях в асимметричной потенциальной яме атомы проводят больше времени: на расстояниях — меньших или больших равновесного значения r_0?
- Поясните, где использовался закон сохранения энергии в приведенном в тексте примере косвенного расчета теплового эффекта химической реакции.
- Опишите различие потенциальных кривых для взаимодействия атомов в молекуле и атомов различных молекул.
- Как связать глубину и положение потенциальной ямы для кривой $U(r)$, описываемой потенциалом (6), со значениями феноменологических параметров a и b?
- Опишите различие представлений о молекулах в физике и в химии.

§ 13. Кристаллы

Атомы большинства веществ, находящихся в твердом состоянии, образуют регулярную *периодическую решетку.* Многие из таких веществ внешне не выглядят как кристаллы, ибо они построены из множества хаотически ориентированных мельчайших кристалликов, т. е. имеют *поликристаллическую структуру.* К ним относятся многие горные породы, металлы и сплавы. Крупные одиночные кристаллы называются

монокристаллами. В природе встречаются монокристаллы различных размеров: от громадных (до сотен килограммов) кристаллов горного

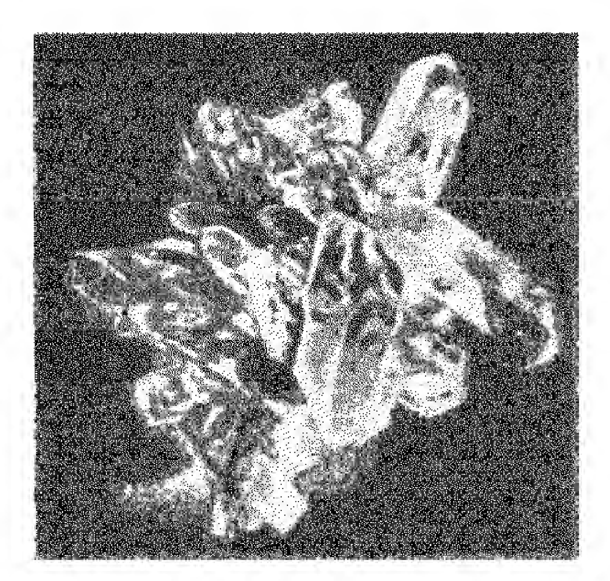


Рис. 38. Друза кристаллов горного хрусталя

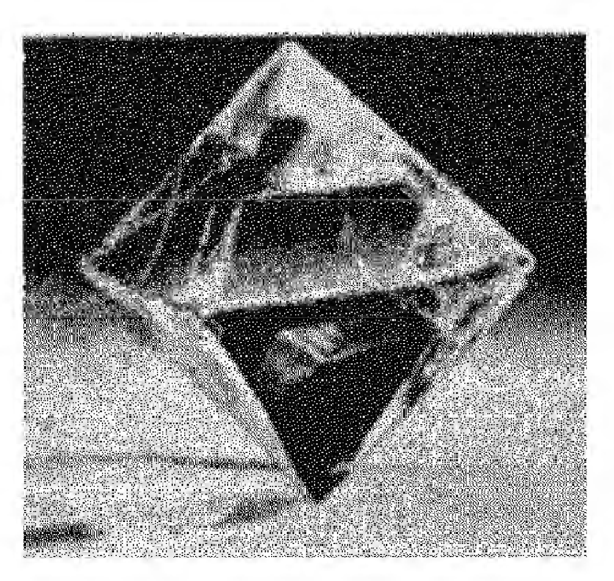


Рис. 39. Кристалл флюорита (CaF_2), имеющий форму октаэдра

хрусталя (кварца, рис. 38), флюорита (рис. 39), полевого шпата до мелких кристаллов алмаза и др.

Монокристаллы и поликристаллы. Для многих практических целей кристаллы выращивают в лабораториях. Это так называемые синтетические кристаллы. Наибольшее значение имеют полупроводниковые, пьезоэлектрические и сегнетоэлектрические кристаллы. Последние годы все большее практическое применение получают так называемые *жидкие кристаллы* — вещества, сочетающие в себе некоторые свойства жидкостей (текучесть) с определенной пространственной упорядоченностью в расположении молекул, проявляющейся в анизотропии ряда физических свойств, характерной для твердых кристаллов.

По характеру сил связи между атомами или молекулами кристаллы делят на четыре основные группы: ионные, ковалентные, металлические и молекулярные кристаллы.

Геометрия кристаллов. Выросшие в равновесных условиях монокристаллы имеют форму многогранников той или иной симметрии. Грани таких кристаллов плоские, ребра между гранями прямолинейные, углы между соответствующими гранями кристаллов одного и того же вещества постоянные. В этих макроскопических свойствах кристаллов находит свое отражение их регулярное внутреннее строение.

Пространственную структуру кристаллов можно представлять себе как совокупность повторяющихся в пространстве одинаковых *элементарных ячеек*. Весь макроскопический образец кристалла построен из таких элементарных ячеек подобно тому, как стена по-

строена из одинаковых кирпичей. Благодаря такому строению для кристаллов характерен *дальний порядок* в расположении атомов, простирающийся на макроскопические расстояния. Этим структура кристаллов отличается от аморфных тел (таких, как смола, пластмассы и т. д.) и жидкостей, где определенный порядок в расположении молекул распространяется только на ближайшее окружение каждой молекулы.

Симметрия кристаллической решетки. Кристаллическая решетка, характеризующая равновесное расположение атомов в идеальном бесконечно протяженном кристалле, обладает определенной симметрией. Прежде всего это *трансляционная симметрия*, т. е. симметрия по отношению к параллельным переносам, при которых бесконечная кристаллическая решетка совмещается сама с собой.

Наименьшее расстояние, при трансляции на которое решетка совмещается сама с собой, называется пространственным *периодом*. Трехмерная пространственная решетка характеризуется тремя основными периодами, которые, очевидно, совпадают с ребрами элементарной ячейки.

Кроме трансляционной симметрии кристаллическая решетка характеризуется некоторой *точечной симметрией*, т. е. преобразованиями симметрии, оставляющими неподвижной какую-либо точку. Только такую симметрию и могут иметь тела конечных размеров. К элементам точечной симметрии относятся центр симметрии, простые оси симметрии и плоскости симметрии. Например, наличие у кристаллической решетки оси симметрии третьего порядка означает, что при повороте вокруг нее на треть оборота (120°) решетка совмещается сама с собой.

Тела конечных размеров могут иметь оси симметрии любого порядка. Например, в основании правильной пирамиды может лежать правильный многоугольник с любым числом сторон. Но у бесконечной кристаллической решетки могут быть оси симметрии только второго, третьего, четвертого и шестого порядков. Связано это с тем, что точечная симметрия здесь должна сочетаться с трансляционной.

Понять, что осей других порядков, кроме перечисленных, быть не может, проще всего на примере покрытия плоскости правильными многоугольниками. Действительно, паркет из одинаковых плиток в виде правильных многоугольников, которыми можно выстлать пол, может состоять только из треугольников, квадратов и шестиугольников. Конечно, плитки могут быть и вытянутыми прямоугольниками, но это даст ось симметрии только второго порядка.

Кристаллические системы и классы. У разных кристаллических решеток по-разному сочетаются упоминавшиеся выше элементы симметрии. Оказывается, что существует 230 различных сочетаний

этих элементов, называемых пространственными группами. Все они были теоретически предсказаны русским кристаллографом Е. С. Федоровым и независимо от него немецким ученым А. Шенфлисом.

230 пространственных групп по характеру симметрии могут быть отнесены к одному из 32 классов, которые группируются в 7 кристаллических систем (или сингоний). К одному кристаллическому классу относятся решетки с одинаковыми наборами элементов точечной симметрии, приводящих к совмещению эквивалентных направлений в кристалле. К одной кристаллической системе относятся решетки с одинаковыми наборами элементов симметрии элементарной ячейки. В частности, кристалл NaCl (см. рис. 37) относится к кубической сингонии и имеет гранецентрированную кубическую решетку. В такой решетке одинаковые ионы (например, ионы Na^+) находятся в вершинах и в центрах граней элементарного куба. Если же одинаковые ионы расположены в вершинах куба и в его центре, то решетка называется объемноцентрированной кубической. Именно так кристаллизуется CsCl.

Симметрия и физические свойства. Все кристаллы одного класса имеют одинаковую симметрию макроскопических физических свойств. Например, кубические кристаллы ведут себя как изотропные тела в отношении прохождения света, электро- и теплопроводности, теплового расширения. Но они анизотропны в отношении упругих и пьезоэлектрических свойств.

Симметрию кристаллической решетки и расположение атомов в элементарной ячейке можно определить с помощью рентгеновского структурного анализа, электронографии и нейтронографии. Во всех этих методах используется явление дифракции волн на трехмерной периодической структуре, которая по существу представляет собой дифракционную решетку. Несмотря на различную физическую природу рентгеновских (т. е. электромагнитных) волн и волн материи (электронных или нейтронных), создаваемая ими дифракционная картина позволяет определить расположение атомов в кристаллической решетке и расстояния между ними, т. е. периоды решетки. Дифракционная картина возникает, когда длина волны меньше периода решетки. В противном случае кристалл ведет себя как сплошная среда, лишенная внутренней структуры.

Молекулярные кристаллы. Остановимся подробнее на характере сил связи между атомами или молекулами в кристаллах. Описанная выше классификация кристаллов по симметрии не зависит от типа этих сил, но для каждого типа сил характерны определенные структуры.

Наиболее простые представители молекулярных кристаллов это кристаллы инертных газов (аргон, криптон, ксенон), хотя построены

они не из молекул, а из отдельных атомов. Такие кристаллы существуют лишь при низких температурах. Атомы инертных газов имеют целиком заполненные внешние электронные оболочки, поэтому пространственное распределение электронного заряда в них сферически-симметрично, а значения энергии ионизации очень высоки.

Энергия связи атомов в молекулярном кристалле мала по сравнению с энергией ионизации (порядка одного процента и менее). Поэтому распределение электронов в атомах кристалла незначительно отличается от распределения в свободных атомах, оставаясь почти сферически-симметричным. Такие атомы можно рассматривать как шары одинакового радиуса, образующие структуру с плотной упаковкой, при которой объем свободного пространства между соприкасающимися шарами минимален.

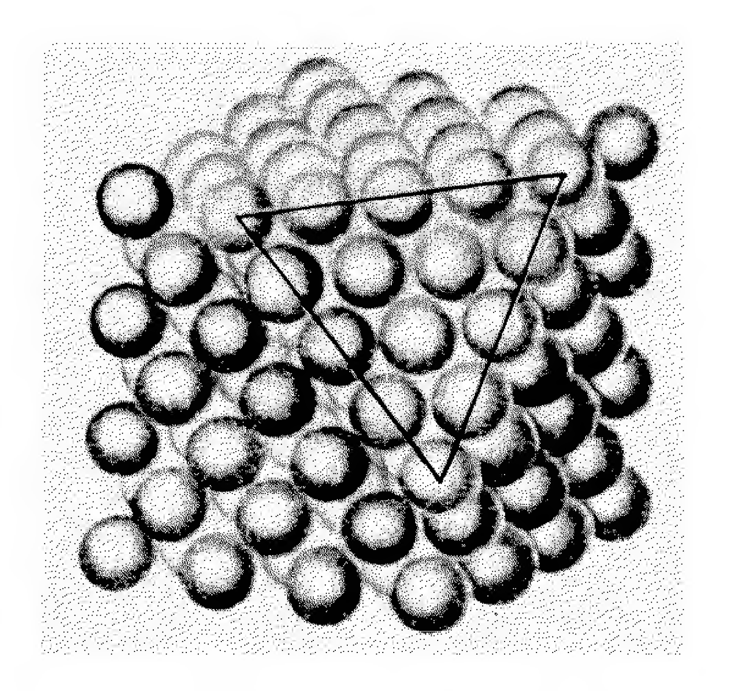


Рис. 40. Гранецентрированная кубическая структура с плотной упаковкой. Плотноупакованные слои твердых шаров лежат в плоскостях, перпендикулярных диагонали куба. Одна такая плоскость видна на месте срезанной вершины куба

Расположить одинаковые твердые шары в пространстве так, чтобы остающийся между ними объем был минимален, можно двумя способами. Один способ приводит к структуре, обладающей гексагональной симметрией. Другой способ дает структуру с кубической симметрией, а именно гранецентрированную кубическую решетку (рис. 40). Такую структуру имеют кристаллы инертных газов. Часть общего объема, занимаемого шарами (коэффициент заполнения), составляет 0,74 как для кубической, так и для гексагональной структур с плотной упаковкой.

Аналогичную структуру имеют молекулярные кристаллы, образованные не отдельными атомами, а одинаковыми неполярными молекулами, такими, как молекула метана CH_4. В молекулярных кристаллах атомы связаны в молекулы значительно сильнее, чем связаны между собой молекулы, образующие кристалл. Поэтому молекулы здесь в значительной степени сохраняют свою индивидуальность, взаимодействуют между собой так же, как и в газообразной фазе, и для описания их взаимодействия можно использовать рассмотренные в предыдущем параграфе модельные потенциалы,

например потенциал Леннарда—Джонса. Самый обширный класс молекулярных кристаллов составляют разнообразные органические соединения. Они могут быть образованы как небольшими молекулами, так и гигантскими (белки и вирусы).

Ионные кристаллы. Ионные кристаллы построены из положительных и отрицательных ионов. Так как в среднем ионы противоположного знака расположены ближе друг к другу, чем ионы одного знака, то кулоновское притяжение преобладает над отталкиванием, и образование кристалла оказывается энергетически выгодным. Электростатическое взаимодействие ионов дает правильное, согласующееся с опытом значение энергии связи кристалла, но одно лишь это взаимодействие не может объяснить устойчивости кристалла: для существования равновесных расстояний между ионами обязательно должны действовать силы отталкивания, которые не сводятся к кулоновскому взаимодействию ионов. Природа этих сил вскрывается только в квантовой механике. Они обусловлены тем, что электроны подчиняются принципу Паули. С такими силами мы уже сталкивались при обсуждении равновесия атомов в молекуле.

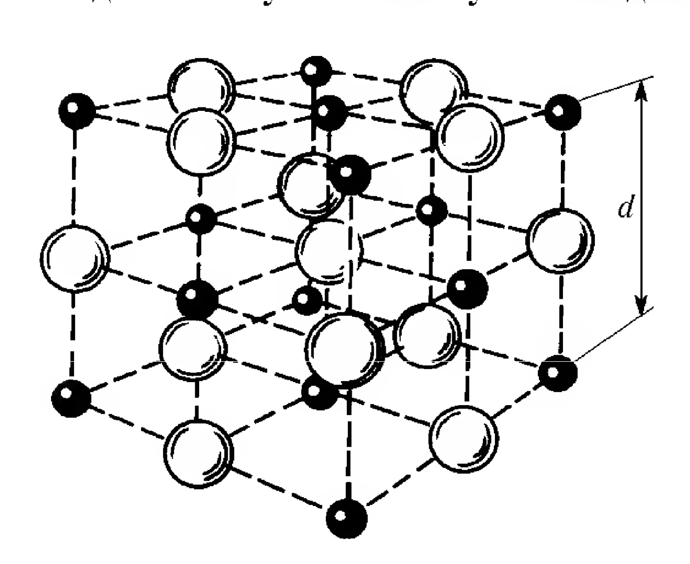


Рис. 41. Модель кристалла NaCl из соприкасающихся твердых шаров двух размеров, которые соответствуют ионам Na^+ и Cl^- и схема их расположения в пространстве

Электронные оболочки всех ионов простого, т. е. состоящего из двух сортов ионов, кристалла до конца заполнены электронами так же, как у атомов инертных газов. Это значит, что распределение заряда в них сферически-симметрично и можно ожидать, что в ионном кристалле эта сферическая симметрия сохраняется, несколько нарушаясь только в области соприкосновения электронных оболочек соседних ионов. Другими словами, в моделях структуры ионных кристаллов, как и в случае кристаллов инертных газов, можно использовать представление о твердых непроницаемых шарах, соприкасающихся друг с другом. Однако здесь в общем случае ионам противоположных знаков соответствуют шары различных размеров, и от того, насколько сильно различаются эти размеры, кристалл образует решетку того или иного типа. Так, например, в уже упоминавшемся ионном кристалле NaCl, схематическое взаимное расположение ионов которого было показано на рис. 37, в дей-

ствительности более мелкие ионы Na^+ располагаются между крупными ионами Cl^-, так что почти не остается пустых промежутков (рис. 41). При этом как те, так и другие ионы образуют одинаковые гранецентрированные кубические решетки, вложенные друг в друга.

Ковалентные кристаллы. В ковалентных кристаллах, таких, как углерод (алмаз), германий и кремний, связь между одинаковыми атомами имеет ту же природу, что и связь атомов в молекуле водорода, где ковалентная связь реализуется в своем простейшем виде. Ковалентная связь весьма прочно скрепляет атомы в кристалле. По своей силе она сравнима со связью в ионных кристаллах, несмотря на то, что осуществляется она между нейтральными атомами, а не между ионами.

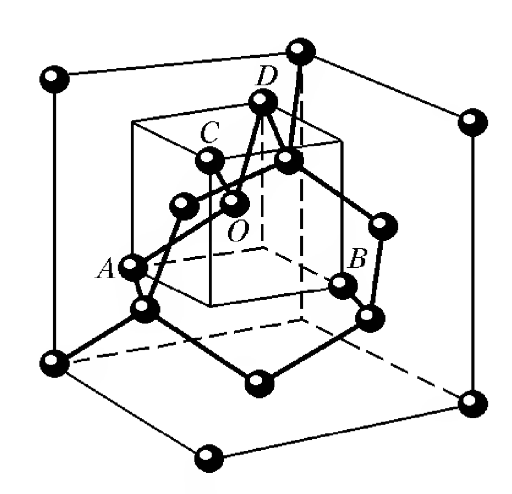


Рис. 42. Пространственное расположение атомов в решетке типа алмаза. Атом *O* находится в центре тетраэдра, образованного его ближайшими соседями — атомами *A*, *B*, *C* и *D*

Ковалентная связь характеризуется явно выраженными свойствами насыщенности и направленности. Так, в кристаллах углерода, германия и кремния, имеющих структуру алмаза, каждый атом помещается в центре тетраэдра, образованного четырьмя такими же атомами, являющимися его ближайшими соседями (рис. 42). Такое расположение приводит к «просторной» в геометрическом смысле упаковке атомов кристалла. Коэффициент заполнения для структуры алмаза равен 0,34, что существенно меньше его значения 0,74 для структуры плотной упаковки. В отличие от плотной упаковки, где число ближайших соседей равно 12, в структуре типа алмаза с тетраэдрическими связями каждый атом имеет только четырех ближайших соседей.

Такой характер структуры ковалентных кристаллов обусловлен тем, что ковалентная связь образуется обычно двумя электронами, по одному от каждого из соединяющихся атомов. Атому углерода, кремния или германия не хватает четырех электронов для образования целиком заполненной внешней оболочки. Поэтому они в состоянии образовать именно четыре направленные связи с соседями в решетке. Этим объясняется направленность и насыщаемость связей в ковалентных кристаллах. Кристаллы с ковалентными и ионными типами связи можно рассматривать как некоторые предельные случаи, между которыми имеется целый ряд кристаллов, обладающих промежуточными типами связи. Сюда относятся многие полупроводниковые материалы (GaAs, InSb, ZnS).

Металлические кристаллы. Металлы образуют еще один тип характерных кристаллических структур. Они отличаются высокой электропроводностью, и поэтому можно думать, что в значительной части электроны в металле должны быть свободными, не связанными с определенными атомами, чтобы иметь возможность перемещаться по кристаллу на макроскопические расстояния. Такие способные принимать участие в электропроводности электроны называются *электронами проводимости.*

Металлическая связь, как правило, возникает между атомами, имеющими один—два электрона на рыхлой внешней оболочке, которые сравнительно слабо связаны с компактным остовом атома, содержащим заполненные электронные оболочки. Сюда прежде всего относятся щелочные и щелочноземельные металлы. Кристаллы щелочных металлов можно представлять себе в виде правильно расположенных положительных ионов, погруженных в более или менее однородную отрицательно заряженную «жидкость» из электронов.

В отличие от ковалентных кристаллов, связи атомов в металлах не обладают свойствами направленности и насыщаемости. Поэтому металлы, как и молекулярные кристаллы инертных газов, имеют тенденцию кристаллизоваться в плотноупакованные и близкие к ним структуры. Например, гексагональную плотную упаковку образуют кристаллы бериллия, кобальта, цинка. Кубическую гранецентрированную решетку с таким же коэффициентом заполнения образуют кристаллы алюминия, меди, золота и кобальта (другой кристаллической модификации). Кубическую объемноцентрированную решетку образуют кристаллы хрома, лития, молибдена и ферромагнитного железа.

Водородная связь. Рассмотренная выше классификация кристаллов по четырем основным типам по характеру сил связи не охватывает всех случаев. В частности, из нее выпадают кристаллы с так называемыми водородными связями. Водородная связь имеет преимущественно ионный характер. В предельном случае такой связи атом водорода теряет свой единственный электрон, отдавая его одному из двух атомов молекулы, и превращается в «голый» протон, имеющий ничтожно малые размеры по сравнению с исходным атомом водорода. Протон может уместиться в узкой «щели» между атомами, осуществляя связь между ними. Именно так обстоит дело в отрицательном ионе HF_2^- (рис. 43).

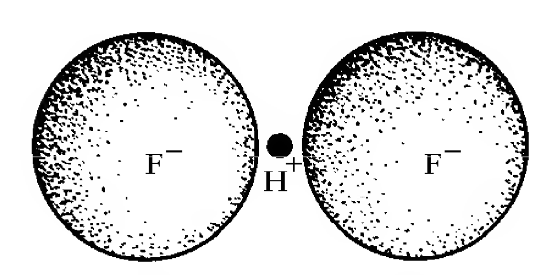


Рис. 43. Схема образования водородной связи между ионами фтора в HF_2^-

Водородная связь ответственна за взаимодействие между молекулами воды (H_2O) и обусловливает вместе с электростатическим притяжением электрических дипольных моментов этих молекул удивительные физические свойства воды и льда.

- Чем отличаются поликристаллы и монокристаллы с точки зрения их строения и физических свойств?
- Что такое элементарная ячейка кристалла?
- В чем заключается трансляционная симметрия кристаллической решетки?
- Почему в кристаллах могут существовать оси симметрии только второго, третьего, четвертого и шестого порядков?
- Покажите, что одинаковыми плитками в виде правильных пятиугольников нельзя настлать паркетный пол. Почему это невозможно в случае плиток в виде семиугольников, восьмиугольников, и т. д.
- Почему для видимого света, т. е. электромагнитных волн с длиной волны от 400 до 700 нм, нельзя наблюдать дифракцию на кристаллической решетке?
- Поясните, каким образом плотно упакованные шары могут образовать гексагональную структуру, учитывая, что центры трех соприкасающихся шаров лежат в вершинах правильного треугольника.
- Почему одно лишь электростатическое взаимодействие не в состоянии объяснить устойчивость решетки ионных кристаллов?
- Какие особенности заполнения электронных оболочек ионов дают возможность рассматривать их в моделях решеток ионных кристаллов как твердые непроницаемые шары?
- Почему для ковалентных кристаллов характерны структуры с малым числом ближайших соседей?
- Чем объясняется существование электронов проводимости в металлических кристаллах?

IV. ОСНОВЫ ТЕРМОДИНАМИКИ

§ 14. Основные понятия термодинамики

Окружающий нас мир, воспринимаемый нами через ощущения, состоит из макроскопических объектов, т. е. тел, которые велики по сравнению с атомными размерами и содержат огромное число атомов и молекул. Этот мир необычайно разнообразен и сложен; он включает в себя газы, плазму, жидкости, твердые тела, биологические организмы разных степеней развития. Изучением этих объектов занимаются разные естественные науки — химия, биология, геология и т. д. Роль физики заключается в установлении наиболее общих, фундаментальных законов, лежащих в основе всего сущего и характеризующих любые макроскопические системы. Главная цель — выяснить каким образом, исходя из небольшого числа твердо установленных простых физических законов, прийти к пониманию наблюдаемого на опыте чрезвычайно разнообразного поведения макроскопических систем.

До сих пор мы изучали простые физические системы, для каждой из которых были справедливы свои законы движения, своя динамика. В одних случаях это законы классической механики, в других — электродинамики, в третьих — квантовой механики. При этом мы, как правило, идеализировали изучаемую систему, отвлекаясь от присущего всем без исключения объектам свойства, а именно от того, что либо сама система, либо ее составные части совершают хаотическое тепловое движение. Закономерности этого движения — общие для объектов любой физической природы, описываемых разными динамическими законами. Универсальный характер теплового движения придает исключительную роль науке, занимающейся изучением его закономерностей.

Типичные макроскопические системы, с которыми мы сталкиваемся в повседневной жизни, содержат порядка 10^{25} взаимодействующих атомов, молекул или других микрообъектов. Нам известны их строение и взаимодействие. Физические законы, описывающие динамическое поведение этих микрообъектов (законы классической и квантовой механики, электродинамики), также хорошо известны. Мы уверены в их справедливости применительно к поведению атомов любой макроскопической системы, будь то жидкость, твердое тело или живой организм. Казалось бы, этого в принципе достаточно для того, чтобы на основе известной микроструктуры макроскопической системы вывести ее наблюдаемые свойства.

Однако оказывается, что реализовать такую программу совершенно невозможно, пока мы не располагаем методами, которые находились бы в соответствии с необычайной сложностью таких систем. Сложность макроскопической системы означает не только невозможность интересоваться поведением всех входящих в нее молекул или атомов. Во многих случаях сложность приводит к появлению новых качеств, которые могут оказаться весьма неожиданными.

Для объяснения свойств макроскопических систем, состоящих из огромного числа частиц, необходимо прежде всего сформулировать новые понятия, отвечающие такому качеству, как необычайная сложность системы. Исторически при изучении макроскопических систем независимо сложились два различных подхода — статистический и термодинамический.

Статистический подход, или *статистическая механика*, основывается на определенных представлениях о строении вещества. Термодинамический подход, или *термодинамика*, представляет собой феноменологическую теорию, основанную на небольшом числе твердо установленных на опыте законов. Исторически термодинамика появилась раньше статистической механики, когда еще не существовало сколько-нибудь надежно подтвержденных на опыте представлений о молекулярном строении вещества.

Центральные понятия термодинамики вводятся не с помощью представлений о внутреннем строении изучаемой системы, а на основе эксперимента. Термодинамика оперирует только макроскопическими величинами: давлением, температурой, энергией системы и т. д.

Термодинамика концентрирует внимание на закономерностях превращений энергии. Она позволяет установить, в каком направлении могут протекать различные физические или химические процессы в тех или иных системах. Термодинамика вскрывает глубокие связи между различными свойствами вещества: располагая экспериментальными данными о каких-то одних свойствах определенного вещества, можно методами термодинамики рассчитать некоторые другие его свойства. В этом разделе будут подробно изложены основы термодинамического подхода.

Термодинамическая система. Одно из основных понятий термодинамики — *термодинамическая система*. Под термодинамической системой понимается совокупность тел любой физической природы и любого химического состава, характеризуемая некоторым числом *макроскопических параметров*. Такими параметрами могут быть, в частности, давление, объем, напряженность электрического и магнитного полей, электрический и магнитный дипольный моменты и т. д. Определение макроскопических параметров в термодинамике дается указанием способа их измерения на опыте.

Например, газ в сосуде характеризуется давлением, которое измеряется так же, как в гидростатике, и объемом, который зависит от размеров сосуда. Диэлектрик, помещенный в электрическое поле, поляризуется, т. е. приобретает электрический дипольный момент, значение которого также рассматривается как макроскопический параметр.

Внешние и внутренние параметры. Если какую-либо совокупность тел принять за интересующую нас термодинамическую систему, то все остальные тела следует рассматривать как внешние по отношению к этой системе. В соответствии с этим и характеризующие систему макроскопические параметры подразделяют на внешние и внутренние. Внешними называют такие параметры, значения которых задаются не входящими в рассматриваемую систему телами. Значения внутренних параметров определяются реакцией самой системы на внешние условия. Некоторые из параметров в зависимости от условий могут выступать либо как внешние, либо как внутренние. Например, давление газа в сосуде может рассматриваться как внутренний параметр, если фиксировать объем газа. Объем в этом случае является внешним параметром. Если же газ находится в сосуде, закрытом поршнем, к которому приложена постоянная сила (рис. 44), то внешним параметром будет давление, так как его значение задается внешними условиями. Объем газа здесь можно рассматривать как внутренний параметр.

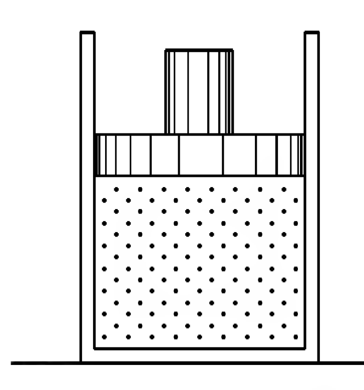


Рис. 44. В сосуде, закрытом поршнем, давление газа задается внешними условиями

Состояние термодинамической системы. В механике состояние системы в некоторый момент времени определяется заданием значений координат и импульсов всех входящих в нее частиц. Понимаемое в таком смысле состояние мы будем называть *микроскопическим состоянием* системы или *микросостоянием*. Наряду с ним можно рассматривать *макроскопическое состояние* или *макросостояние*, характеризуемое заданием только макроскопических параметров. Одному и тому же макросостоянию системы может соответствовать множество ее различных микросостояний. В термодинамике рассматриваются только макроскопические состояния, которые в дальнейшем для краткости будем называть просто состояниями.

Термодинамический процесс. Всякое изменение в термодинамической системе, связанное с изменением хотя бы одного из ее макроскопических параметров, называется *термодинамическим процессом*. Другими словами, термодинамический процесс в системе есть изменение ее состояния.

Стационарные и равновесные состояния. Состояние называется *стационарным*, если все макроскопические параметры системы не меняются со временем. Стационарное состояние может поддерживаться внешними по отношению к системе процессами. Например, неизменный во времени перепад температур между концами стержня можно создать, нагревая все время один его конец и охлаждая другой.

Если стационарность состояния не обусловлена внешними процессами, то состояние называется *равновесным* или состоянием *термодинамического равновесия*. Когда макроскопическая система находится в состоянии термодинамического равновесия, то все макроскопические части, на которые можно мысленно или реально разбить эту систему, также находятся в равновесии как сами по себе, так и друг с другом.

Время релаксации и локальное равновесие. Из любого неравновесного состояния *замкнутая*, или *изолированная, система*, т. е. система, никак не взаимодействующая с окружением, приходит в состояние термодинамического равновесия за некоторое время, называемое *временем релаксации*. Некоторые из макроскопических параметров имеют определенные для всей системы значения только в состоянии равновесия.

В некоторых случаях всю систему можно мысленно разделить на слабо взаимодействующие между собой макроскопические части, равновесие в которых устанавливается значительно быстрее, чем во всей системе. При этом можно говорить о *локальном равновесии*, когда макроскопические параметры имеют определенные значения для каждой из таких частей, но эти значения могут быть разными для различных частей системы. Характерные времена установления локального равновесия могут существенно (на несколько порядков) различаться для разных макроскопических параметров.

В неравновесных системах существуют *потоки*, связанные с переносом массы, заряда, энергии и т. п. из одного места в другое. Когда в системе существует локальное равновесие, эти потоки обычно пропорциональны градиентам (т. е. быстроте изменения в пространстве) определенных макроскопических параметров. Например, поток вещества пропорционален градиенту давления, поток электрического заряда — электрический ток — пропорционален градиенту потенциала, т. е. напряженности электрического поля, и т. д.

Закон сохранения и превращения энергии в термодинамике. Как уже отмечалось, термодинамика занимается поисками соотношений между различными свойствами вещества, не углубляясь в детали его внутреннего строения. Она основана на изучении энергетических превращений и опирается на закон сохранения энергии.

Закон сохранения энергии является всеобъемлющим законом природы. Мы не знаем ни одного исключения из него. Более того, если обнаруживается, что в каком-либо физическом явлении энергия не сохраняется, то это означает, что здесь просто оказалась не учтенной какая-то из ее форм. Именно многообразие различных видов энергии и особенности превращений энергии из одних видов в другие и обусловливают эффективность термодинамического подхода к изучению природы. Обилие существующих форм энергии, способных к взаимным превращениям, выделяет закон сохранения энергии из других фундаментальных законов, таких как законы сохранения импульса и момента импульса.

Существование механической, электромагнитной, ядерной и других форм энергии означает, что этой физической величине нельзя дать общего определения, применимого сразу ко всем ее видам. В то же время отсюда следует, что общий закон сохранения энергии нельзя вывести из других физических законов, хотя для каждого из видов энергии ее сохранение вытекает из соответствующих динамических законов для тех явлений, в которых не происходят превращения энергии в другие виды.

Великий французский ученый А. Пуанкаре выразил это следующими словами: «Поскольку мы не в состоянии дать общее определение энергии, закон сохранения энергии следует рассматривать просто как указание на то, что существует нечто, остающееся постоянным в любом физическом процессе. К каким бы открытиям ни привели нас будущие эксперименты, мы заранее знаем, что и тогда будет нечто, обладающее способностью сохраняться, и это нечто мы можем называть *энергией*».

Внутренняя энергия. Все тела состоят из атомов и молекул, находящихся в непрестанном тепловом, хаотическом движении. Поэтому даже если тело в целом неподвижно и имеет нулевую потенциальную энергию, оно тем не менее обладает энергией, связанной с внутренним движением составляющих его атомов и молекул и с их взаимодействием. Иначе говоря, любому коллективу атомов или молекул всегда присуща некоторая *внутренняя энергия*. Если рассматривать внутреннюю структуру вещества, то всегда можно конкретизировать формы и носителя этой внутренней энергии. Но термодинамика сознательно отвлекается от внутренней структуры изучаемых систем, поэтому в термодинамике внутреннюю энергию необходимо рассматривать как *особую форму энергии*. При этом, разумеется, полезно помнить, что в конечном счете внутренняя энергия всегда связана с хаотическим тепловым движением образующих систему частиц и их взаимодействием друг с другом. Поэтому иногда внутреннюю энергию называют тепловой энергией.

Внутренняя энергия зависит от интенсивности теплового движения. Для характеристики этой интенсивности служит особая физическая величина — *температура*. Температура играет важную роль не только в термодинамике, но и вообще в физике.

Температура. Представление о температуре, как и представление о силе, вошло в науку через наши чувственные восприятия. Наши ощущения позволяют различать качественные градации степени нагретости: холодный, теплый, горячий. Однако в основу количественного определения температуры и построения температурной шкалы должны быть положены объективные физические закономерности, свободные от субъективизма чувственных восприятий. В феноменологическом подходе к тепловым явлениям температура вводится через понятие термодинамического равновесия. К понятию температуры можно прийти на основе конкретных примеров и последующего обобщения.

Если два тела, температуры которых при оценке на основе наших ощущений сильно различаются (например, раскаленный кусок металла и холодная вода), привести в соприкосновение, то горячее тело будет охлаждаться, а холодное — нагреваться до тех пор, пока в этой системе не прекратятся всякие макроскопические изменения. В таком случае говорят, что эти тела пришли в состояние термодинамического равновесия и имеют *одинаковую температуру*. Опыт показывает, что в состояние равновесия в конце концов приходит любое число соприкасающихся тел. Если какие-то два тела находятся в равновесии с третьим телом, то они находятся в равновесии между собой.

Поэтому для суждения о равенстве или различии температур двух тел A и B необязательно приводить их в соприкосновение друг с другом. Можно воспользоваться вспомогательным третьим телом C, приводимым поочередно в контакт с телами A и B. Таким способом можно убедиться только в равенстве или в различии температур тел A и B. Для того чтобы изменить их температуры, нужно установить *температурную шкалу*, т. е. по определению ввести правила, по которым каждой температуре ставится в соответствие определенное число. В выборе таких правил имеется большой произвол, необходимо лишь обеспечить взаимную однозначность температур и сопоставляемых им чисел.

Измерение температуры. Термометр. При создании температурной шкалы нужно основываться на каком-либо свойстве вещества, зависящем от температуры. Опыт показывает, что практически все физические свойства тел изменяются с температурой. Так, при нагревании многие тела расширяются, электрическое сопротивление металлов возрастает с повышением температуры, а полупроводников убывает, и т. д. Все подобные явления в принципе можно использовать для создания приборов для измерения температуры — *термометров*.

Основной частью термометра является *термометрическое тело*, приводимое в тепловой контакт с объектом, температуру которого надо измерить. В жидкостных термометрах термометрическим телом служит либо ртуть, либо подкрашенный спирт, а о температуре судят по изменению объема жидкости — по высоте столбика в капилляре. В термометрах сопротивления термометрическим телом служит, например, металлическая проволока, а температура определяется по ее электрическому сопротивлению.

Важные требования, предъявляемые к термометру, — чувствительность и точность измерений, а также воспроизводимость результатов. Для практического применения важно, чтобы термометр не изменял температуру тела, с которым он приводится в контакт. Поэтому его чувствительный элемент должен быть достаточно мал. Наконец, важно и то, насколько быстро устанавливается термодинамическое равновесие термометра с этим телом.

Эмпирическая температурная шкала. Устройство большинства термометров основано на предположении, что положенное в основу измерения физическое свойство термометрического тела линейно зависит от температуры. Для построения шкалы выбираются две так называемые *реперные точки*, которым приписываются произвольные значения температуры, а шкала между ними делится на равные части. Этим устанавливается единица измерения температуры.

Раньше в качестве реперных точек выбирались точка плавления льда и точка кипения воды при нормальном атмосферном давлении. В шкале Цельсия первой реперной точке присваивалась температура 0, а второй — 100 градусов Цельсия (0°C—100°C). Современная температурная шкала, сохраняя преемственность со старой, основывается на так называемой тройной точке воды, которая обладает гораздо лучшей воспроизводимостью.

Установленная описанным выше способом шкала температуры называется *эмпирической температурной шкалой*, а измеряемая температура — эмпирической температурой. В зависимости от выбора термометрического тела можно осуществить множество эмпирических температурных шкал. Из всех эмпирических шкал наибольшими преимуществами обладает так называемая идеально-газовая шкала температуры, в наименьшей степени зависящая от химической природы вещества, выбранного в качестве термометрического тела. Выбирая именно такую шкалу, мы приближаемся к идеалу, когда температурная шкала вообще не зависит от термометрического тела. Принципиальная возможность построения термодинамической («абсолютной»), не зависящей от термометрического тела шкалы температуры, устанавливается в термодинамике. Этот вопрос будет рассмотрен ниже.

Высокие и низкие температуры. На практике для измерения температур в разных интервалах, в том числе очень высоких и очень

низких температур, используют методы, основанные на разных физических явлениях. Например, для измерения очень высоких температур (выше 1000 °C) применяют оптические пирометры, проградуированные на основе законов теплового излучения. Очень низкие температуры (ниже 1 K) определяют по измерениям магнитной восприимчивости некоторых парамагнитных солей. Разумеется, для измерения температур в таких областях приходится использовать свои реперные точки. Важно, чтобы используемые в разных интервалах температур методы измерений давали совпадающие результаты в тех областях, где они перекрываются.

Изолированная система и термодинамическое равновесие. Как уже отмечалось, система, не обменивающаяся энергией и веществом с внешними телами, называется изолированной. В термодинамике на основе обобщения опытных данных принимается как постулат следующее утверждение: любая изолированная система рано или поздно обязательно приходит в состояние термодинамического равновесия и самопроизвольно из него выйти не может.

Опыт показывает, что в состоянии равновесия не все макроскопические параметры, которые можно использовать для описания системы, являются независимыми. Независимы только внешние параметры. Например, если некоторое количество газа находится в сосуде определенного объема при некоторой температуре, то его давление имеет вполне определенное значение, которое однозначно выражается через объем и температуру газа. Другими словами, для определенного количества газа давление является функцией объема и температуры.

Это положение имеет универсальный характер. Как обобщение опытных данных в термодинамике принимается, что в состоянии равновесия в любой системе значения внутренних параметров являются функциями внешних параметров и температуры.

Уравнение состояния. Функциональная связь, выражающая зависимость какого-либо внутреннего параметра от значений внешних параметров и температуры, называется *уравнением состояния*. Конкретный вид этого уравнения зависит от рассматриваемой термодинамической системы. Для каждого вещества характер этой функциональной связи индивидуален. Поэтому термодинамические свойства описываются своим для каждого вещества уравнением состояния. В термодинамике постулируется только существование уравнения состояния, а его конкретный вид берется из опыта.

Оказывается, что некоторую информацию о свойствах системы можно получить, опираясь только на факт существования уравнения состояния и не зная его явного вида. Если же известен явный вид уравнения, можно получить более детальную информацию. В следующем параграфе будет рассмотрено уравнение состояния для

наиболее простой термодинамической системы газа, находящегося в закрытом сосуде.

- Что понимают под термодинамической системой? Какими параметрами она характеризуется?
- Что такое внешние и внутренние параметры термодинамической системы?
- В чем различие понятий состояния системы в механике и термодинамике?
- В каком случае состояние термодинамической системы является равновесным? Чем оно отличается от стационарного состояния?
- Перечислите известные вам виды энергии и приведите примеры явлений, в которых происходят превращения энергии из одного вида в другой.
- Почему в термодинамике внутреннюю энергию следует рассматривать как особый вид энергии? Что представляет собой внутренняя энергия с точки зрения микроструктуры вещества?
- Приведите примеры, иллюстрирующие субъективный характер оценки температуры воздуха, воды или окружающих предметов на основе чувственных восприятий.
- Как вводится понятие температуры в термодинамике?
- Какие требования предъявляются к термометрическому телу, выбранному для эмпирической температурной шкалы? Почему для измерения температуры в интервале от 1°C до 10°C нельзя использовать воду в качестве термометрического тела?
- Почему ртутный и спиртовой термометры с равномерными шкалами дают совпадающие показания, строго говоря, только в реперных точках?
- В каком направлении развиваются процессы в замкнутых термодинамических системах?
- Что такое уравнение состояния в термодинамике? Выводится ли оно теоретически в рамках термодинамики?

§ 15. Уравнение состояния газа

Возьмем некоторое количество газа определенного химического состава, например азота, кислорода или воздуха, и заключим его в сосуд, объем которого можно изменять по своему усмотрению. Будем считать, что у нас имеется манометр, т. е. прибор для измерения давления газа, и термометр для измерения его температуры. Опыт показывает, что перечисленные макроскопические параметры полностью характеризуют газ как термодинамическую систему в том случае, когда этот газ состоит из нейтральных молекул, не обладающих собственным дипольным моментом.

В состоянии термодинамического равновесия не все эти параметры независимы, они связаны между собой уравнением состояния. Чтобы получить это уравнение, нужно воспользоваться установленными на опыте закономерностями поведения газа при изменении каких-либо внешних параметров.

Газ в сосуде — простая термодинамическая система. Примем сначала, что ни количество газа, ни его химический состав во время опыта не меняются, так что речь пойдет только о трех макроскопических параметрах — давлении p, объеме V и температуре t. Для установления связывающих эти параметры закономерностей удобно зафиксировать значение одного из параметров и следить за изменениями двух других. Будем считать, что вызываемые нами изменения в газе происходят настолько медленно, что в любой момент времени макроскопические параметры p, V, t, характеризующие весь газ в состоянии термодинамического равновесия, имеют вполне определенные значения.

Изопроцессы. Как уже отмечалось, из любого неравновесного состояния термодинамическая система приходит в состояние равновесия за некоторое время — время релаксации. Чтобы при происходящих в системе изменениях макроскопические параметры имели вполне определенные значения, характерное время этих изменений должно быть много больше времени релаксации. Это условие накладывает ограничения на допустимую скорость процесса в газе, при котором сохраняют смысл его макроскопические параметры.

Процессы, протекающие при неизменном значении одного из параметров, принято называть *изопроцессами*. Так, процесс, происходящий при постоянной температуре, называется изотермическим, при постоянном объеме — *изохорическим* (изохорным), при постоянном давлении — *изобарическим* (изобарным).

Закон Бойля–Мариотта. Исторически первым в газе был экспериментально изучен изотермический процесс. Английский физик Р. Бойль и независимо от него французский физик Э. Мариотт установили закон изменения объема при изменении давления: для данного количества любого газа при неизменной температуре объем обратно пропорционален давлению. Обычно закон Бойля—Мариотта записывают в виде

$$pV = \text{const} \qquad (t = \text{const}). \tag{1}$$

Для поддержания постоянной температуры исследуемый газ должен находиться в хорошем тепловом контакте с окружающей средой, имеющей неизменную температуру. В этом случае говорят, что газ находится в контакте с *термостатом* — большим тепловым резервуаром, на состояние которого не влияют любые изменения, происходящие с исследуемым газом.

Закон Бойля—Мариотта хорошо выполняется для всех газов и их смесей в широком диапазоне температур и давлений. Отклонения от этого закона становятся существенными лишь при давлениях, в несколько сотен раз превышающих атмосферное, и при достаточно низких температурах.

Проверить справедливость закона Бойля—Мариотта можно совсем простыми средствами. Для этого достаточно иметь запаянную с одного конца стеклянную трубку, в которой столбик ртути закрывает некоторое количество воздуха (трубка Мельде). Объем воздуха можно измерять линейкой по длине воздушного столба в трубке (рис. 45), а о давлении можно судить по высоте столбика ртути при разных ориентациях трубки в поле тяжести.

Для наглядного изображения изменений состояния газа и происходящих с ним процессов удобно использовать так называемые p—V-диаграммы, где по оси абсцисс откладываются значения объема, а по оси ординат — давления. Кривая на p—V-диаграмме, соответствующая изотермическому процессу, называется *изотермой*. Как следует из закона Бойля—Мариотта, газовые изотермы представляют собой гиперболы (рис. 46). Чем выше температура, тем дальше от координатных осей расположена соответствующая изотерма.

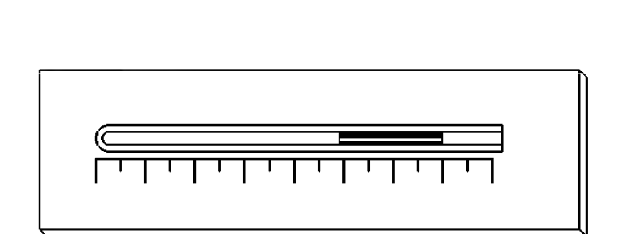


Рис. 45. Простейший прибор для проверки закона Бойля—Мариотта (трубка Мельде)

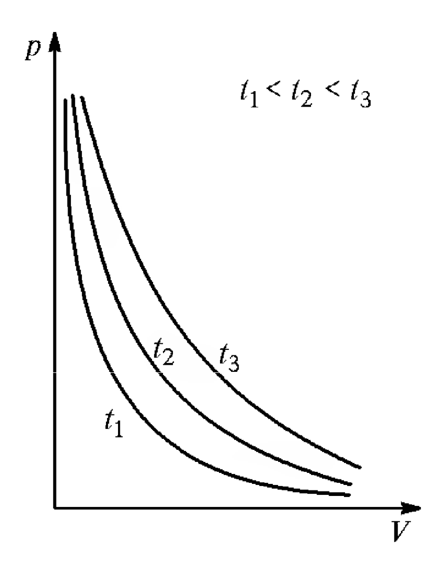


Рис. 46. Изотермы газа на p—V-диаграмме

Закон Шарля. Зависимость давления газа от температуры при неизменном объеме была экспериментально установлена французским физиком Ж. Шарлем. Согласно закону Шарля, давление газа при постоянном объеме линейно зависит от температуры:

$$p = p_0(1 + \alpha t^\circ) \quad (V = \text{const}), \qquad (2)$$

где p_0 — давление газа при 0 °C. Оказывается, что температурный коэффициент давления α одинаков для всех газов и равен 1/273 °C^{-1}.

Закон Гей-Люссака. Аналогичный вид имеет и зависимость объема газа от температуры при неизменном давлении. Это было установлено на опыте французским физиком Ж. Гей-Люссаком, который нашел,

что температурный коэффициент расширения одинаков для всех газов. Значение этого коэффициента оказалось таким же, как и коэффициента α в законе Шарля. Таким образом, закон Гей-Люссака можно записать в виде

$$V = V_0(1 + \alpha t) \quad (p = \text{const}), \tag{3}$$

где V_0 — объем газа при 0 °C.

Совпадение температурных коэффициентов в законах Шарля и Гей-Люссака не случайно и свидетельствует о том, что эти устанавливаемые на опыте газовые законы не являются независимыми. Ниже мы подробнее остановимся на этом.

Газовый термометр. Тот факт, что выражаемая законами Шарля и Гей-Люссака зависимость давления или объема от температуры одинакова для всех газов, делает особенно удобным выбор газа в качестве термометрического тела. Хотя на практике использовать газовые термометры в силу их громоздкости и тепловой инерционности неудобно, именно по ним производится градуировка других термометров, более удобных для практических применений.

Шкала Кельвина. Зависимость давления или объема от температуры в законах Шарля и Гей-Люссака станет еще проще, если перейти к новой температурной шкале, потребовав, чтобы линейная зависимость превратилась в прямую пропорциональность.

Изобразив выражаемую формулой (3) зависимость объема газа от температуры (рис. 47) и продолжив график $V(t)$ влево до пересечения с осью температуры, легко убедиться, что продолжение графика пересекает ось t при значении температуры, равном −273 °C, поскольку $\alpha = 1/273\ °\text{C}^{-1}$. Именно в эту точку нужно поместить начало новой температурной шкалы, чтобы можно было записать уравнения (2) и (3) как прямую пропорциональность. Эту точку называют абсолютным нулем температуры. Масштаб новой шкалы, т. е. единица измерения температуры, выбирается так же, как и в шкале Цельсия. На новой температурной шкале нулю градусов Цельсия соответствует температура $T_0 = +273$ градуса (точнее 273,15), а любая другая температура T связана с соответствующей температурой t по шкале Цельсия соотношением

$$T = t + T_0, \quad \text{или} \quad T = t + 273. \tag{4}$$

Введенная здесь температурная шкала называется *шкалой Кельвина*, а единица измерения, совпадающая с градусом шкалы Цельсия, называется *кельвином* и обозначается буквой К. Иногда эта шкала называется Международной практической шкалой температуры.

При использовании температурной шкалы Кельвина график закона Гей-Люссака принимает вид, показанный на рис. 48, а форму-

лы (2) и (3) можно записать в виде

$$p = p_0 \frac{T}{T_0} \quad (V = \text{const}), \tag{5}$$

$$V = V_0 \frac{T}{T_0} \quad (p = \text{const}). \tag{6}$$

Коэффициент пропорциональности V_0/T_0 в (6) характеризует наклон графика $V(T)$ на рис. 48.

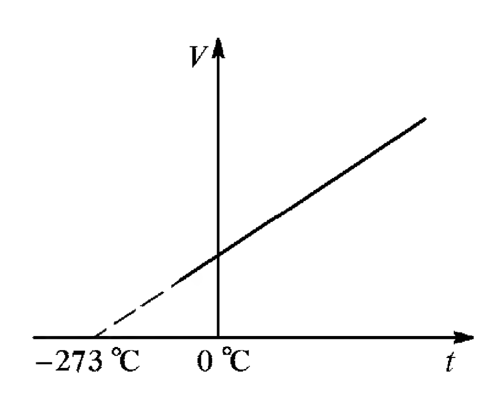


Рис. 47. Выражаемая законом Гей-Люссака зависимость объема газа от температуры при постоянном давлении

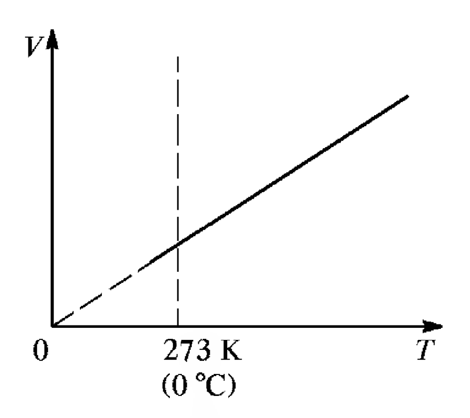


Рис. 48. График закона Гей-Люссака в температурной шкале Кельвина

Уравнение состояния газа. Экспериментальные газовые законы дают возможность установить уравнение состояния газа. Для этого достаточно воспользоваться любыми двумя из приведенных законов. Пусть некоторое количество газа находится в состоянии с объемом V_1, давлением p_1 и температурой T_1. Переведем его в другое (промежуточное) состояние, характеризуемое тем же значением температуры $T' = T_1$ и некоторыми новыми значениями объема V' и давления p'. При изотермическом процессе выполняется закон Бойля-Мариотта, поэтому

$$p_1 V_1 = p' V'. \tag{7}$$

Теперь переведем газ из промежуточного состояния в конечное состояние с тем же значением объема $V_2 = V'$, что и в промежуточном состоянии, и некоторыми значениями давления и температуры p_2 и T_2. При изохорическом процессе выполняется закон Шарля, поэтому

$$p_2 = p' \frac{T_2}{T'} = p' \frac{T_2}{T_1}, \tag{8}$$

поскольку $T' = T_1$. Подставляя в (8) $p' = p_1 V_1 / V'$ из (7) и учитывая, что $V' = V_2$, окончательно получаем

$$\frac{p_1 V_1}{T_1} = \frac{p_2 V_2}{T_2}. \tag{9}$$

Мы изменили все три макроскопических параметра p, V и T, и тем не менее соотношение (9) показывает, что для данного количества газа (числа молей $\nu = m/M$) комбинация параметров pV/T имеет одно и то же значение, в каком бы состоянии этот газ не находился. Это означает, что уравнение (9) представляет собой уравнение состояния газа. Его называют уравнением Клапейрона.

В приведенном выводе уравнения (9) не использовался закон Гей-Люссака. Однако легко видеть, что в нем содержатся все три газовых закона. Действительно, полагая в (9) $p_1 = p_2$, получаем для изобарического процесса соотношение $\frac{V}{T} = \text{const}$, что соответствует закону Гей-Люссака.

Уравнение Менделеева–Клапейрона. Возьмем один моль газа при нормальных условиях, т. е. при $T_0 = 273{,}15$ К (0 °С) и нормальном атмосферном давлении $p_0 = 101325$ Па (1 атм). В соответствии с установленным на опыте законом Авогадро один моль любого газа (гелия, азота, кислорода и т. д.) занимает при нормальных условиях одинаковый объем $V_0 = 22{,}414$ литра. Поэтому для одного моля любого газа комбинация p_0V_0/T_0, обозначаемая через R и называемая *универсальной газовой постоянной* (или молярной газовой постоянной), имеет одно и то же значение:

$$R = \frac{p_0V_0}{T_0} = \frac{101325 \cdot 22{,}414 \cdot 10^{-3}}{273{,}25} \, \frac{\text{Па} \cdot \text{м}^{-3}}{\text{К} \cdot \text{моль}} = 8{,}3145 \, \frac{\text{Дж}}{\text{моль} \cdot \text{К}}. \qquad (10)$$

С учетом (10) уравнение состояния одного моля любого газа можно записать в виде

$$pV = RT. \qquad (11)$$

Уравнение (11) легко обобщить для произвольного количества газа. Так как при тех же значениях температуры и давления ν молей газа занимают в ν раз больший объем, чем 1 моль, то

$$pV = \nu RT. \qquad (12)$$

В таком виде уравнение состояния газа впервые было получено русским ученым Д. И. Менделеевым. Поэтому его называют уравнением Менделеева—Клапейрона.

Идеальный газ. Уравнение состояния газа (11) или (12) было получено на основе установленных на опыте газовых законов. Эти законы выполняются приближенно: условия их применимости различны для разных газов. Например, для гелия они справедливы в более широком диапазоне температур и давлений, чем для углекислого газа. Приближенным является и уравнение состояния, полученное из приближенных газовых законов.

Введем в рассмотрение физическую модель — *идеальный газ*. Под этим будем понимать систему, для которой уравнение (11) или (12) является точным. Замечательной особенностью идеального газа является то, что его внутренняя энергия пропорциональна абсолютной температуре и не зависит от объема, занимаемого газом.

Как и во всех других случаях использования физических моделей, применимость модели идеального газа к тому или иному реальному газу зависит не только от свойств самого газа, но и от характера вопроса, на который требуется найти ответ. Такая модель не позволяет описать особенности поведения различных газов, но выявляет свойства, общие для всех газов.

С применением уравнения состояния идеального газа можно познакомиться на примере конкретных задач.

Задачи

1. В одном баллоне объемом V_1 находится азот при давлении p_1. В другом баллоне объемом V_2 находится кислород при давлении p_2. Температура газов совпадает с температурой окружающей среды. Какое установится давление газов, если открыть кран трубки, соединяющей эти баллоны между собой?

Решение. После открывания крана газ из баллона с более высоким давлением будет поступать в другой баллон. В конце концов давление в баллонах выровняется, а газы перемешаются. Даже если в процессе перетекания газов температура изменилась, после установления теплового равновесия она снова сравняется с температурой окружающего воздуха.

Для решения задачи можно воспользоваться уравнением состояния идеального газа. Обозначив через ν_1 и ν_2 количество газов в баллонах до открывания крана, имеем

$$p_1V_1 = \nu_1 RT, \qquad p_2V_2 = \nu_2 RT. \tag{13}$$

В конечном состоянии смесь газов содержит $\nu_1 + \nu_2$ молей, занимает объем $V_1 + V_2$ и находится при давлении p, которое нужно определить. Применяя к смеси газов уравнение Менделеева—Клапейрона, имеем

$$p(V_1 + V_2) = (\nu_1 + \nu_2)RT. \tag{14}$$

Выражая ν_1 и ν_2 из уравнений (13) и подставляя в (14), находим

$$p = \frac{p_1V_1 + p_2V_2}{V_1 + V_2}. \tag{15}$$

В частном случае, когда исходные давления газов одинаковы, давление смеси после установления равновесия остается таким же. Интересен предельный случай $V_2 \to \infty$, соответствующий замене второго сосуда атмосферой. Из (15) при этом получаем $p = p_2$, где p_2 — давление атмосферы. Такой результат очевиден из общих соображений.

Обратим внимание на то, что выражаемый формулой (15) результат соответствует тому, что давление смеси газов равно сумме парциальных давлений каждого из газов, т. е. давлений, которые имел бы каждый из газов, занимая при той же температуре весь объем. Действительно, парциальные давления

p_1' и p_2' каждого газа можно найти с помощью закона Бойля—Мариотта:

$$p_1' = p_1 \frac{V_1}{V_1+V_2}, \quad p_2' = p_2 \frac{V_2}{V_1+V_2}.$$

Видно, что полное давление p, равное сумме парциальных давлений $p_1' + p_2'$, выражается формулой (15). Утверждение, что давление смеси химически невзаимодействующих газов равно сумме парциальных давлений, называется законом Дальтона.

2. Истопив печь, в дачном домике температуру воздуха повысили от 0 до 20 °С. Как при этом изменилась плотность воздуха?

Решение. Ясно, что объем помещения при протапливании печи не изменился, так как тепловым расширением стен можно пренебречь. Если бы мы нагревали воздух при неизменном объеме V в закрытом сосуде, его давление возросло бы, но плотность осталась бы неизменной. Но дачный домик не герметичен, поэтому неизменным остается давление p воздуха, равное наружному атмосферному давлению. Ясно, что при повышении температуры T должна измениться масса m воздуха в помещении: какая-то его часть должна выйти через щели наружу.

Подсчитать плотность воздуха проще всего, основываясь на уравнении состояния:

$$pV = \frac{m}{M} RT, \tag{16}$$

где $M = 29$ г/моль — молярная масса воздуха. Отсюда для плотности воздуха $\rho = m/V$ имеем

$$\rho = \frac{Mp}{RT}. \tag{17}$$

При неизменном давлении плотность обратно пропорциональна абсолютной температуре. Поэтому для отношения плотностей ρ/ρ_0 при температурах $T = 293$ К и $T_0 = 273$ К соответственно находим

$$\frac{\rho}{\rho_0} = \frac{T_0}{T} = \frac{273}{293} = 0{,}93.$$

3. В бутылку емкостью $V_0 = 0{,}5$ л сквозь герметичную пробку вставлена открытая с обоих концов тонкая стеклянная трубка диаметром $d = 5$ мм. В трубке находится столбик подкрашенной воды. Как будет изменяться положение столбика с температурой? Оцените, на сколько градусов нужно нагреть бутылку, чтобы столбик воды переместился на 1 см.

Решение. При расширении воздуха в бутылке его давление, равное сумме атмосферного давления и гидростатического давления столбика воды в трубке, остается неизменным, пока вода не выливается из трубки. Поэтому, в соответствии с законом Гей-Люссака, объем воздуха в бутылке пропорционален абсолютной температуре: $V \sim T$. Отсюда следует, что $\Delta V \sim \Delta T$ и

$$\frac{\Delta V}{V} = \frac{\Delta T}{T}. \tag{18}$$

В данном случае $\Delta V = S\Delta h$, где $S = \pi d^2/4$ — внутреннее сечение трубки, а Δh — изменение положения столбика воды при нагревании воздуха. Из (18) следует, что

$$\Delta h = \frac{V}{S}\frac{\Delta T}{T}. \tag{19}$$

Поскольку внутренний объем трубки много меньше объема бутылки, то при расчетах можно в (19) заменить объем воздуха V на емкость бутылки V_0.

Ясно, что столбик воды не будет вытолкнут из трубки только при очень малых изменениях температуры. Чтобы оценить изменение температуры, при котором столбик поднимается на заданное расстояние Δh, перепишем (19) следующим образом:

$$\Delta T = T \frac{S}{V_0} \Delta h.$$

Полагая для оценки $T = 300$ К, получаем $\Delta T = 0{,}1$ К. Приведенная оценка показывает, что с помощью этого очень простого устройства можно обнаружить изменение температуры вплоть до 0,01 К, так как легко заметить изменение положения столбика на 1 мм.

- Что такое время релаксации для термодинамической системы?
- Какие ограничения должны быть наложены на скорость протекания процессов в газе, чтобы в любой момент времени имели смысл макроскопические параметры p, V, t описывающие газ в состоянии равновесия?
- Чем определяется числовое значение константы в правой части уравнения закона Бойля—Мариотта (1)?
- Что имеют в виду, когда говорят, что изучаемая система находится в контакте с термостатом?
- Предложите способ проверки закона Бойля—Мариотта с помощью описанного в тексте прибора (см. рис. 45).
- Какие преимущества дает выбор газа в качестве термометрического тела?
- Как связан выбор начала отсчета температур в шкале Кельвина со значением температурного коэффициента расширения газа?
- Как устанавливается связь температур, измеренных по шкале Цельсия и шкале Кельвина?
- Выведите уравнения состояния газа, используя законы Бойля—Мариотта и Гей-Люссака.
- Уравнение Клапейрона было получено с использованием только двух газовых законов, однако содержит в себе все три закона. Как это связано с тем фактом, что у газов температурные коэффициенты давления и объема одинаковы?
- Что такое универсальная газовая постоянная? Как она связана с законом Авогадро?
- Какую физическую систему называют идеальным газом? Чем определяются условия применимости этой модели? От чего зависит внутренняя энергия идеального газа?
- Можно ли объяснить установленный на опыте закон Дальтона для смеси газов, опираясь на уравнение Менделеева—Клапейрона?
- Оцените массу воздуха, выходящего из щелей отапливаемого помещения объемом 50 м^3 при повышении температуры от 0 до 20 °С.
- Как изменится чувствительность к изменениям температур простого устройства, описанного в задаче 3, если верхнее отверстие трубки заткнуть?

§ 16. Первый закон термодинамики

В состоянии равновесия термодинамическая система характеризуется определенными значениями макроскопических параметров. Эти параметры связаны между собой уравнением состояния.

Функция состояния. Любой внутренний параметр является функцией внешних параметров и температуры. В этом случае говорят, что он представляет собой *функцию состояния*. Функциями состояния можно считать любые характеризующие систему физические величины, значения которых не зависят от предыстории и полностью определяются состоянием системы в данный момент времени.

Внутренняя энергия как функция состояния. Одной из важнейших функций состояния термодинамической системы является ее внутренняя энергия, природа которой в термодинамике не конкретизируется. Внутренняя энергия рассматривается в термодинамике как особая форма энергии, способная к превращениям в другие формы, например, в механическую — кинетическую или потенциальную — энергию. Когда система оказывается в данном состоянии, ее внутренняя энергия принимает соответствующее этому состоянию значение независимо от того, из какого другого состояния и каким способом система перешла в данное состояние. Это характерное для данного состояния значение внутренней энергии зависит только от внешних параметров и от температуры.

Зависимость внутренней энергии от температуры у всех встречающихся в окружающем мире термодинамических систем такова, что с ростом температуры внутренняя энергия также увеличивается.

Работа и теплота как формы изменения внутренней энергии. При взаимодействии термодинамической системы с окружающей средой происходит обмен энергией. При этом можно выделить два принципиально различных способа передачи энергии от системы к внешним телам (или наоборот). Первый способ изменения внутренней энергии связан с *совершением работы* действующими на систему внешними силами. Здесь работа понимается в том же смысле, как в механике и электродинамике. В термодинамике *работа* выражается через макроскопические параметры, например как произведение давления на изменение объема.

Второй способ изменения внутренней энергии, не связанный с совершением работы, называется *теплопередачей*. При первом способе изменение внутренней энергии равно работе внешних сил над системой. При втором способе полученное системой количество внутренней энергии называется *теплотой*.

Подчеркнем, что понятие работы пришло в термодинамику из механики и электродинамики. В отличие от этого понятие теплоты возникает в самой термодинамике и не имеет аналогов в механике и электродинамике. Важно, что и работа, и теплота — это физические величины, характеризующие *процессы*, в которых изменяется внутренняя энергия термодинамической системы, в то время как внутренняя энергия характеризует ее *состояние*. Поэтому имеет физический смысл понятие энергии, запасенной системой, но нельзя говорить, что система запасла какое-то количество работы или теплоты. Или, другими словами, можно говорить о совершенной работе и переданной теплоте, но не имеет смысла говорить о содержащейся в системе работе или теплоте. Это значит, что ни теплота, ни работа не являются, в отличие от внутренней энергии, функциями состояния.

Работа в термодинамике. Рассмотрим подробнее некоторые вопросы, связанные с процессами совершения работы и теплообмена. Получим формулу, позволяющую рассчитать работу, совершаемую при медленном изменении объема какой-либо термодинамической системы, например газа, находящегося в сосуде, закрытом подвижным поршнем (рис. 49). Пусть приложенная к поршню внешняя сила F медленно перемещает поршень из положения *1* в близкое положение *2* так, чтобы силу F можно было считать постоянной. Совершаемая силой F работа A равна произведению силы на перемещение Δx:

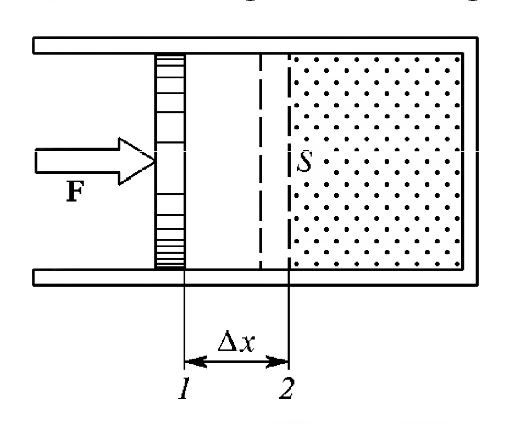


Рис. 49. К вычислению работы при изменении объема

$$A = F \cdot \Delta x. \qquad (1)$$

Если движение поршня происходит без ускорения, то сила F уравновешивается силой давления газа на поршень:

$$F = pS, \qquad (2)$$

где S — площадь поршня. При подстановке F из (2) в выражение (1) возникает произведение $S \cdot \Delta x$, характеризующее изменение ΔV занимаемого газом объема V. Так как объем V уменьшается, то

$$S \cdot \Delta x = -\Delta V. \qquad (3)$$

Теперь выражение (1) для работы действующей на поршень внешней силы принимает вид

$$A = -p\Delta V. \qquad (4)$$

Когда внешняя сила сжимает газ, т. е. $\Delta V < 0$, ее работа положительна. И наоборот, когда газ расширяется, медленно передвигая

поршень, работа удерживающей поршень внешней силы F отрицательна ($A < 0$).

Наряду с работой внешней силы можно рассматривать работу силы, с которой газ давит на поршень. Эту работу принято называть работой, совершаемой газом, или *работой газа*. Очевидно, что работа газа A' отличается от работы A внешней силы F только знаком:

$$A' = -A. \tag{5}$$

Адиабатические процессы. Если сжатие или расширение газа происходит без теплообмена с окружающей средой, то изменение ΔU его внутренней энергии U равно работе внешней силы:

$$\Delta U = A. \tag{6}$$

Учитывая соотношение (5), это выражение можно записать и таким образом:

$$A' = -\Delta U. \tag{7}$$

Когда газ расширяется без теплообмена с окружением, он совершает положительную работу $A' > 0$ за счет своей внутренней энергии U, запас которой при этом убывает ($\Delta U < 0$).

Процессы, происходящие в отсутствие теплообмена с окружающей средой, называются адиабатическими (адиабатными). Для практической реализации таких процессов нужны определенные условия: либо заключить систему в оболочку с низкой теплопроводностью, либо проводить процесс достаточно быстро, чтобы внутренняя энергия не успела заметным образом измениться из-за теплообмена. В этом случае говорят об адиабатической изоляции системы.

Работа внешних электрических сил. Внешние силы, совершающие работу над системой, могут иметь электромагнитную природу. Если в системе есть свободные электрические заряды, то при приложении внешнего электрического поля силы поля совершают работу, когда происходит перемещение этих зарядов. Например, если заряд q перемещается из точки с потенциалом φ_1 в точку с потенциалом φ_2, то силы поля совершают работу

$$A = q(\varphi_1 - \varphi_2) = -q \cdot \Delta\varphi, \tag{8}$$

где $\Delta\varphi$ — разность потенциалов между конечной и начальной точками. Работа внешних электрических сил в отсутствие теплообмена связана с изменением внутренней энергии системы тем же соотношением (6). Здесь также в некоторых случаях удобно говорить не о работе внешних сил над системой, а о работе самой системы над внешними телами. Так, при зарядке аккумулятора естественно говорить о работе внешних сил, за счет которой происходит увеличе-

ние внутренней энергии. Однако при разрядке аккумулятора, когда он используется в качестве источника питания в какой-либо электрической цепи, естественно говорить о работе, совершаемой самим аккумулятором, а не внешними силами, хотя эти величины, как и прежде, отличаются только знаком.

Теплота. Перейдем к более подробному обсуждению теплообмена. Чтобы он происходил, нужно привести систему в контакт с телами, имеющими другую температуру. Передача внутренней энергии при теплообмене происходит от системы в окружающую среду, если температура системы выше, чем у окружения, и наоборот, система получает энергию, если ее температура ниже, чем у окружения. Процесс теплопередачи происходит до тех пор, пока существует разность температур. Если в процессе теплопередачи не совершается работа, то изменение внутренней энергии равно переданному системе количеству теплоты Q:

$$\Delta U = Q. \qquad (9)$$

Процесс теплопередачи может происходить не только при непосредственном соприкосновении тел, но и через излучение. Именно таким путем Земля получает энергию Солнца.

В той или иной мере теплообмен происходит практически во всех макроскопических процессах. Представление об адиабатическом процессе — это некоторая идеализация.

Первый закон термодинамики. В середине XIX века исследованиями многих ученых, в первую очередь немецкого врача Р. Майера, английского физика Дж. Джоуля и немецкого физика Г. Гельмгольца была твердо установлена эквивалентность теплоты и работы, т. е. возможность их сравнения и измерения в одних и тех же единицах. Тем самым была заложена основа для обобщения закона сохранения энергии на тепловые процессы и для формирования общего понятия энергии как физической величины.

Согласно закону сохранения энергии энергия не возникает и не исчезает: она только переходит из одной формы в другую или от одной физической системы к другой. Первый закон термодинамики математически выражает количественную сторону закона сохранения и превращения энергии. Его можно сформулировать следующим образом.

Изменение внутренней энергии системы $\Delta U = U_2 - U_1$ в результате процесса перехода из начального состояния в конечное равно сумме совершаемой над системой внешними силами работы A и полученного системой количества теплоты Q:

$$\Delta U = A + Q. \qquad (10)$$

Содержание первого закона термодинамики можно выразить и таким образом:

$$Q = \Delta U + A' \tag{11}$$

— сообщенная системе теплота равна сумме изменения внутренней энергии и совершенной системой работы. Естественно, что эти формулировки эквивалентны, так как $A' = -A$ в силу соотношения (5).

Приведенные формулировки первого закона равнозначны утверждению о невозможности вечного двигателя, т. е. машины, совершающей работу без затрат энергии.

Квазистатические процессы на p–V-диаграмме. Процесс перехода системы из данного начального состояния в определенное конечное состояние можно осуществить различными способами. При этом совершаемая над системой работа A и передаваемая ей теплота Q зависят от способа перехода. Разумеется, их сумма, равная изменению энергии системы, во всех случаях будет одна и та же.

Поясним это на примере простейшей термодинамической системы — идеального газа. Равновесное состояние одного моля идеального газа характеризуется заданием двух параметров, например давления и объема. Третий параметр — температура находится при этом из уравнения состояния Менделеева—Клапейрона. Поэтому *состояние* моля идеального газа можно изобразить *точкой* на двумерной диаграмме (рис. 50), а *квазистатический процесс*, представляющий собой последовательность равновесных состояний, — непрерывной *линией*. Работа $\Delta A'$, совершаемая газом при изменении его объема на малую величину ΔV, равна произведению давления на изменение объема:

$$\Delta A' = p \cdot \Delta V. \tag{12}$$

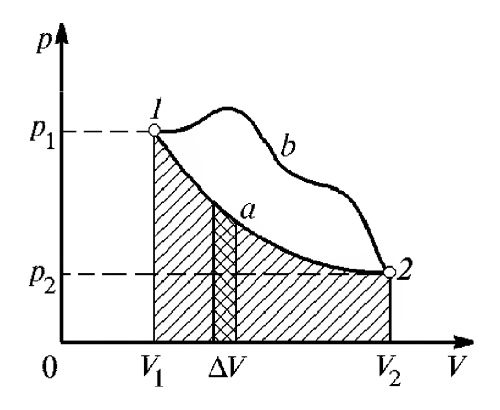


Рис. 50. p—V-диаграмма идеального газа

Из рис. 50 видно, что эта работа численно равна площади заштрихованной полоски высотой p и шириной ΔV. Поэтому работа, совершаемая газом при расширении из состояния *1* в состояние *2* по пути *a*, равна площади криволинейной трапеции, ограниченной графиком процесса и осью абсцисс. Работа, совершаемая внешними силами над газом при этом процессе, как уже отмечалось, отличается от работы газа только знаком. Теперь легко видеть, что работы внешних сил, совершаемые при расширении газа из состояния *1* в состояние *2* по путям *a* и *b*, различны. Но изменение внутренней энергии газа $U_2 - U_1$ не зависит от пути перехода. Следовательно, сообщаемые

газу при этих процессах количества теплоты, как следует из (10), будут также различны.

Теплоемкость. Получаемое системой в ходе квазистатического процесса количество теплоты характеризуется физической величиной — *теплоемкостью.* Теплоемкостью называется отношение переданной системе на участке процесса теплоты ΔQ к происшедшему на этом участке изменению температуры системы ΔT:

$$C = \frac{\Delta Q}{\Delta T}. \tag{13}$$

Количество теплоты, получаемое системой при изменении ее температуры на ΔT, будет неодинаковым для различных процессов. Поэтому будет разной и теплоемкость. Таким образом, теплоемкость является характеристикой не самого вещества, а определенного процесса с данным веществом.

С помощью первого закона термодинамики можно получить выражение для теплоемкости при разных процессах. Для этого удобно воспользоваться первым законом в форме (11). Применяя (11) к малому участку процесса и рассматривая системы, у которых выражение для совершаемой работы имеет вид (12), получим

$$C = \frac{\Delta U}{\Delta T} + p\,\frac{\Delta V}{\Delta T}. \tag{14}$$

Изохорический процесс. Вычислим с помощью (14) теплоемкости одного моля идеального газа при различных процессах. Для изохорического процесса, при котором объем газа остается неизменным, $\Delta V = 0$, и с помощью формулы (14) находим

$$C_V = \frac{\Delta U}{\Delta T}. \tag{15}$$

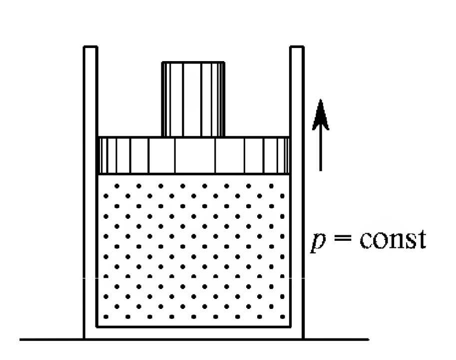


Рис. 51. Изобарическое нагревание газа

Теплоемкость C_V получила название молярной теплоемкости при постоянном объеме. Она характеризует процесс нагревания газа в закрытом сосуде.

Изобарический процесс. Когда газ заключен в цилиндр с вертикальными стенками, закрытый сверху поршнем с грузом, то нагревание газа будет происходить при постоянном давлении (рис. 51). Такое изобарическое нагревание сопровождается расширением газа, при котором он, поднимая поршень с грузом, совершает работу. Поэтому теплоемкость газа при постоянном давлении C_p больше, чем теплоемкость при постоянном объеме C_V. Из уравнения состояния одного

моля идеального газа $pV = RT$ при постоянном давлении p находим $p \cdot \Delta V = R \cdot \Delta T$. Теперь с помощью (14) получаем

$$C_p = C_V + R. \tag{16}$$

Соотношение (16) называется уравнением Майера.

Адиабатический процесс. При отсутствии теплообмена с окружающими телами теплоемкость равна нулю, а первый закон термодинамики принимает вид

$$\Delta U = A, \tag{17}$$

что, естественно, совпадает с (6): изменение внутренней энергии системы происходит только за счет работы внешних сил. Это же самое можно выразить иначе: при адиабатическом процессе совершаемая системой работа A' равна убыли внутренней энергии системы:

$$A' = -\Delta U = U_1 - U_2. \tag{18}$$

При расширении система совершает положительную работу $A' > 0$, но $A < 0$, и ее внутренняя энергия убывает. Для идеального газа, у которого внутренняя энергия растет с увеличением температуры и не зависит от объема, это означает, что при адиабатическом расширении газ будет охлаждаться. При адиабатическом сжатии такой газ будет нагреваться.

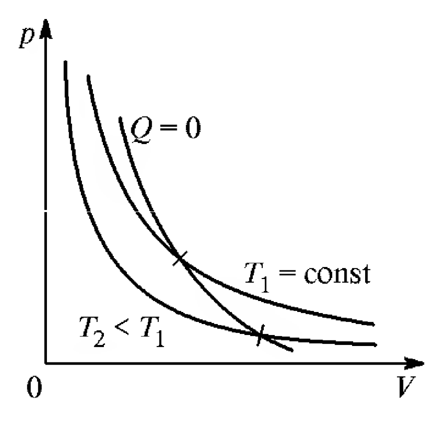


Рис. 52. Изотермы и адиабата идеального газа на p—V-диаграмме

Такое поведение позволяет понять, почему на p—V-диаграмме идеального газа график квазистатического адиабатического процесса идет круче графика изотермического процесса (рис. 52). Действительно, при изотермическом расширении температура постоянна, а при адиабатическом — убывает; поэтому адиабата должна пересекать изотермы, соответствующие все более и более низким температурам, в результате чего давление при расширении падает быстрее, чем на изотерме.

Изотермический процесс. При изотермическом расширении идеальный газ совершает работу только за счет подводимой к нему теплоты, так что его внутренняя энергия во время такого процесса не меняется:

$$A' = Q, \qquad \Delta U = 0. \tag{19}$$

При изотермическом сжатии идеальный газ совершает отрицательную работу. При этом теплота от него отводится: $Q < 0$.

Тепловой двигатель. Тепловыми двигателями называют машины, которые преобразуют выделяющуюся при сжигании топлива внутреннюю энергию в механическую. Первые разные по конструкции и назначению паровые машины появились на рубеже XVII—XVIII веков в Англии и Франции. Спустя полстолетия И. Ползуновым в России и Д. Уаттом в Англии были созданы универсальные паровые машины, пригодные для различных нужд. Их появление произвело настоящую техническую революцию, что дало основание образно назвать последовавшее за этим столетие «веком пара».

Во второй половине XIX века появились двигатели внутреннего сгорания, получившие особенно большое распространение. Наиболее совершенными оказались так называемые четырехтактные двигатели, разработанные немецким конструктором И. Отто.

Применим первый закон термодинамики к анализу работы тепловой машины. Несмотря на большое разнообразие, все тепловые двигатели имеют общий принцип действия. Назначение любого периодически действующего теплового двигателя состоит в совершении механической работы за счет использования внутренней энергии, о чем кратко, но не совсем точно говорят как о превращении теплоты в работу.

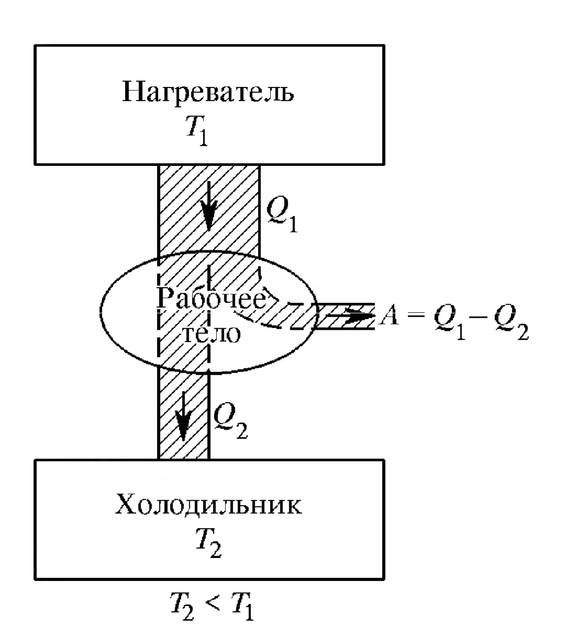


Рис. 53. Принципиальная схема тепловой машины

Принципиальная схема любой тепловой машины, если отвлечься от ее конструктивных особенностей, выглядит так, как показано на рис. 53. Ее обязательными элементами являются два тепловых резервуара: *нагреватель* с некоторой температурой T_1 и *холодильник* с температурой T_2, меньшей температуры нагревателя. Роль холодильника может выполнять окружающая среда. Если просто привести нагреватель в тепловой контакт с холодильником, то внутренняя энергия нагревателя будет передаваться холодильнику путем теплопередачи без совершения работы. Для совершения механической работы обязательно должно быть промежуточное звено — так называемое *рабочее тело*, в качестве которого может быть использован, например, газ в цилиндре, закрытом поршнем. У периодически действующей машины все процессы с рабочим телом повторяются, так что оно периодически возвращается в исходное состояние.

Цикл тепловой машины. Механическая работа совершается при расширении газа в цилиндре, при котором он проходит через ряд состояний вдоль некоторой кривой a на рис. 54. Для того чтобы рабочее тело вернулось в исходное состояние 1, газ в цилиндре нужно сжать, для чего над ним придется совершить работу. Но эта работа обязательно должна быть меньше, чем работа, совершенная газом при расширении. На p—V-диаграмме работа газа равна площади криволинейной трапеции под кривой a. Чтобы при сжатии газа совершилась меньшая работа, кривая сжатия b должна лежать ниже кривой a: сжатие газа должно происходить при более низких температурах, чем расширение.

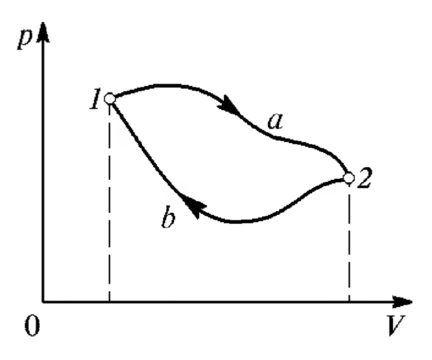


Рис. 54. Цикл теплового двигателя на p—V-диаграмме

Разность площадей под кривыми a и b, т. е. площадь, ограниченная замкнутой кривой 1—a—2—a—1, равна работе A', совершаемой рабочим телом за цикл. На основании первого закона термодинамики эта работа равна разности между теплотой Q_1, полученной рабочим телом за цикл от нагревателя, и теплотой Q_2, отданной им холодильнику, ибо рабочее тело после совершения цикла возвращается в исходное состояние, так что его внутренняя энергия принимает исходное значение:

$$A' = Q_1 - Q_2. \tag{19}$$

КПД тепловой машины. Отношение совершенной за цикл работы A' к количеству теплоты Q_1, полученному за цикл от нагревателя, называется *коэффициентом полезного действия* тепловой машины η:

$$\eta = \frac{A'}{Q_1}. \tag{20}$$

Максимально выгодной была бы тепловая машина, у которой вся полученная теплота Q_1 превращалась бы в работу: $A' = Q_1$. Однако на основании опыта установлено, что это невозможно.

К обсуждению КПД тепловой машины мы еще вернемся при изучении второго закона термодинамики.

- Что такое функция состояния термодинамической системы? Какие физические величины, характеризующие систему, можно рассматривать как функции состояния?
- Поясните, почему внутренняя энергия термодинамической системы является функцией состояния.
- Что такое теплота и работа? Какова их связь с внутренней энергией? Поясните сходство и различие этих физических величин.

- Можно ли говорить о количестве теплоты или работы, запасенной системой? Являются ли теплота и работа функциями состояния системы?
- Пусть на рис. 49 слева от поршня цилиндр сообщается с атмосферой. Что в этом случае следует понимать под силой F, работа которой равна $-p \cdot \Delta V$? Чему равна эта сила, если давление газа в цилиндре равно p, а давление атмосферы равно p_0?
- Поясните, почему работа газа при изменении его объема отличается знаком от работы внешней силы, действующей на поршень. Почему это утверждение справедливо только при достаточно медленном равномерном движении поршня?
- Пусть газ в условиях адиабатической изоляции совершает отрицательную работу. Что при этом происходит с его внутренней энергией?
- Покажите, что первый закон термодинамики в форме (10) или (11) эквивалентен утверждению о невозможности вечного двигателя.
- Поясните, почему точка на $p-V$-диаграмме полностью характеризует состояние определенного количества идеального газа? Чему соответствует любая кривая на такой диаграмме?
- Как по графику некоторого процесса на $p-V$-диаграмме можно определить совершаемую системой работу? В каком случае работа положительна? отрицательна?
- Почему теплоемкость некоторого количества вещества зависит от происходящего с ним процесса? Объясните, почему теплоемкость при постоянном давлении C_p всегда больше теплоемкости при постоянном объеме C_V.
- Чему равно значение теплоемкости для адиабатического процесса? Можно ли провести с идеальным газом такой процесс, при котором теплоемкость будет отрицательной?
- Объяснитс, почсму на $p-V$-диаграммс идсального газа адиабата идст круче изотермы?
- Объясните устройство и принцип действия теплового двигателя.
- Поясните, почему работа, совершаемая тепловым двигателем за один цикл, изображается на $p-V$-диаграмме площадью, охватываемой замкнутой кривой этого цикла.
- Что такое КПД теплового двигателя? Чем определяется его значение?

§ 17. Примеры применения первого закона термодинамики

Первый закон термодинамики, как и закон сохранения энергии в механике, часто дает возможность исследовать тепловые процессы в макроскопических системах даже в тех случаях, когда нам не известны детали микроскопической картины изучаемых явлений.

Энергетический баланс. Первый закон универсален, он применим ко всем без исключения тепловым процессам в любых системах. Как и всякий закон сохранения, он не дает детальной информации о ходе процесса, но позволяет составить уравнение баланса, если заранее известно, какие энергетические превращения происходят в рассматриваемой системе.

В этом параграфе мы рассмотрим примеры использования первого закона термодинамики.

Теплота и внутренняя энергия. Прежде всего сделаем несколько замечаний о смысле входящих в уравнение первого закона величин. Количество переданной теплоты было определено как мера изменения внутренней энергии системы при теплопередаче. Но не всегда подведение к системе теплоты приводит к изменению ее внутренней энергии. Например, при изотермическом расширении идеального газа подведение теплоты не сопровождается увеличением внутренней энергии газа. Внутренняя энергия идеального газа зависит только от температуры и при изотермическом процессе не меняется, но газ совершает работу, и величина этой работы равна подводимому к системе количеству теплоты.

Теплота и работа. Совершение внешними силами механической работы над системой также может не сопровождаться изменением ее внутренней энергии. Если сжимать идеальный газ, принимая меры к тому, чтобы его температура при этом не увеличивалась, то внутренняя энергия газа останется без изменения, а окружающим телам перейдет некоторое количество теплоты, равное совершенной над газом при его сжатии работе.

Пример применения первого закона термодинамики. Применяя первый закон термодинамики, нужно всегда внимательно следить за тем, к каким изменениям в самой системе может привести подведение к ней теплоты и совершение работы. Поясним это на следующем примере. Представьте себе, что в комнате на некоторое время включили электрический нагреватель, в результате чего температура воздуха увеличилась от T_1 до T_2. Может показаться, что в результате этого внутренняя энергия воздуха в комнате увеличилась. Проверим, так ли это. Будем считать воздух идеальным газом. В состоянии теплового равновесия внутренняя энергия одного моля пропорциональна абсолютной температуре, а энергия всего воздуха в комнате пропорциональна количеству (числу молей) газа, находящегося в комнате. Поэтому выражение для энергии можно записать в виде

$$U = C_V \frac{m}{M} T, \qquad (1)$$

где m — масса воздуха, M — его молярная масса, C_V — молярная теплоемкость при постоянном объеме.

Воспользовавшись уравнением Менделеева—Клапейрона $pV = (m/M)RT$, замечаем, что выражению для внутренней энергии воздуха U можно придать вид

$$U = C_V \frac{pV}{R}. \tag{2}$$

Объем V комнаты не изменяется при работе нагревателя, не изменяется и давление p воздуха в комнате. Оно равно атмосферному, поскольку комната не герметична. Тогда из формулы (2) видно, что внутренняя энергия воздуха в комнате не изменяется при протапливании. Сразу возникают вопросы: что происходит с потребляемой от сети электроэнергией и зачем мы вообще включаем электронагреватель? Отвечая на второй вопрос, отметим, что для человека имеет значение не энергия воздуха, а его температура, которая повышается при протапливании. Что касается энергии, потребляемой нагревателем от сети, то она целиком «выходит» наружу: масса воздуха в комнате при нагревании при постоянном давлении уменьшается.

Итак, внутренняя энергия воздуха в комнате при работе нагревателя не меняется, несмотря на повышение температуры. Но никакого парадокса в этом нет. Первый закон термодинамики, разумеется, справедлив и в этом случае. Просто при нагревании изменяется сама рассматриваемая система: количество воздуха в комнате уменьшается.

Еще один пример применения первого закона термодинамики. Рассмотрим теперь стационарный поток идеального газа, протекающего через длинную спиральную трубку — змеевик (рис. 55). На входе змеевика поддерживаются постоянные давление p_1 и температура T_1. На выходе змеевика поддерживаетсяостоянное давление p_2 и измеряется температура T_2. Прежде всего подсчитаем работу, совершаемую при прохождении через змеевик одного моля газа.

О какой работе идет речь? Мы знаем, что газ, а точнее, сила давления газа, совершает работу при перемещении поршня или любого другого тела, ограничивающего занимаемый газом объем. Таким телом является, например, оболочка резинового шарика, деформирующаяся при его наполнении газом. Общее выражение для работы силы давления газа дается формулой (12) предыдущего параграфа:

$$\Delta A' = p \cdot \Delta V. \tag{3}$$

Отметим, что наличие тела, ограничивающего объем газа, совершенно необходимо для того, чтобы можно было говорить о совершаемой газом работе: газ, расширяющийся в пустоту, работы не совершает!

Поясним это. Представим себе сосуд, разделенный перегородкой на две части: по одну сторону от перегородки находится газ, по другую — вакуум. При внезапном удалении перегородки газ заполняет

весь сосуд, объем газа увеличивается ($\Delta V > 0$), давление газа $p > 0$, однако работы газ не совершает: $\Delta A' = 0$. Формула (3) в этом случае неприменима.

Наряду с работой, совершаемой газом, можно рассматривать работу внешних сил, совершаемую над газом при перемещении поршня. Очевидно, что при равномерном перемещении поршня работа ΔA, совершаемая над газом, противоположна по знаку и равна по модулю работе газа $\Delta A'$, т. е. $\Delta A = -p \cdot \Delta V$. Так о какой же работе идет речь? Ведь никаких поршней здесь нет! Однако на самом деле «поршни», т. е. внешние тела, поддерживающие заданные давления на входе и выходе змеевика, здесь есть, и они-то и совершают работу над протекающим газом. Эту работу можно подсчитать следующим образом.

Рассмотрим газ, заключенный между входным S_1 и выходным S_2 сечениями змеевика (рис. 55). Вследствие стационарности потока масса этого газа не меняется со временем, т. е. если за время Δt через входное сечение прошла масса газа Δm, то точно такая же масса вышла из змеевика.

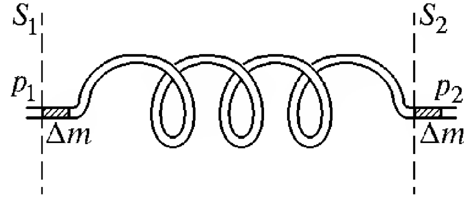


Рис. 55. Прохождение газа через змеевик

Пусть объем, занимаемый массой Δm на входе при давлении p_1 и температуре T_1, равен ΔV_1, а объем на выходе из змеевика равен ΔV_2. Газ, находившийся между сечениями S_1 и S_2, теперь занимает новое положение: идущая за ним порция газа действовала на него с силой $F_1 = p_1 S_1$, идущий впереди газ оказывал сопротивление $F_2 = p_2 S_2$. Таким образом, работа внешних сил над рассматриваемым газом $\Delta A = p_1 \Delta V_1 - p_2 \Delta V_2$. Поскольку во время процесса давления газа на входе и выходе остаются постоянными, то работа внешних сил при проталкивании через змеевик одного моля газа равна

$$A = p_1 V_1 - p_2 V_2, \tag{4}$$

где V_1 и V_2 — объемы, занимаемые одним молем газа при условиях, существующих на входе и на выходе змеевика соответственно. Согласно уравнению Менделеева—Клапейрона можно написать

$$p_1 V_1 = RT_1, \qquad p_2 V_2 = RT_2.$$

Поэтому совершаемая над одним молем газа работа равна

$$A = R(T_1 - T_2). \tag{5}$$

Теперь обратим внимание на одну особенность рассматриваемого примера: в то время как температура газа на входе в змеевик T_1 *задается*, температура на выходе T_2 *измеряется*. Напомним, что дав-

ление в обоих случаях задается. От чего же зависит температура газа на выходе? Очевидно, от условий его прохождения через змеевик. Чем же могут различаться эти условия? Только интенсивностью теплообмена протекающего по змеевику газа с окружающей средой. Рассмотрим вначале случай, когда теплообмен вообще отсутствует, — змеевик адиабатически изолирован от окружающей среды. Какова при этом температура газа T_2 на выходе? Используя первый закон термодинамики

$$Q + A = \Delta U, \tag{6}$$

получаем $A = \Delta U$, так как $Q = 0$. Поскольку внутренняя энергия идеального газа пропорциональна термодинамической температуре, имеем

$$\Delta U = U_2 - U_1 = C_V(T_2 - T_1). \tag{7}$$

Подставляя выражения (5) и (7) в уравнение (6) и учитывая, что в отсутствие теплообмена $Q = 0$, получаем

$$R(T_1 - T_2) = C_V(T_2 - T_1),$$

а так как $C_V > 0$, то отсюда немедленно следует $T_1 = T_2$.

Таким образом, температура идеального газа при адиабатическом прохождении через змеевик не меняется, а совершаемая при этом работа, как видно из (5), равна нулю. Для того чтобы работа была отлична от нуля, необходим теплообмен. Легко убедиться, что, когда газ получает теплоту ($Q > 0$), совершаемая над ним работа отрицательна и, наоборот, при $Q < 0$ совершаемая над газом работа положительна: $A > 0$. Действительно, подставляя в уравнение первого закона (6) выражения (5) и (7), получаем

$$Q = (C_V + R)(T_2 - T_1). \tag{8}$$

Сумма $C_V + R$ равна молярной теплоемкости идеального газа при постоянном давлении C_p, поэтому формулу (8) можно записать и так:

$$Q = C_p(T_2 - T_1). \tag{9}$$

Из этого выражения видно, что знак $\Delta T = T_2 - T_1$ совпадает со знаком Q. Если, например, $Q < 0$, т. е. газ при прохождении через змеевик отдает теплоту, то $T_2 < T_1$, и из выражения (5) видно, что совершаемая при этом над газом работа положительна: $A > 0$.

Змеевик как тепловая машина. Исходя из полученных результатов, попытаемся представить себе, как происходит протекание газа через змеевик. Если в змеевике газ охлаждается, т. е. отдает теплоту ($Q < 0$), то совершаемая над газом работа положительна — внешние силы «проталкивают» газ через змеевик. Если теплота подводится к газу ($Q > 0$), то наш змеевик подобен тепловой маши-

не — газ сам совершает работу над внешними телами. И обратите внимание, что этот результат не зависит от того, какова величина давления газа на выходе и на входе. Единственное условие при этом — давление на входе должно быть больше давления на выходе, иначе газ просто потечет в обратную сторону.

Как было выяснено, при адиабатическом протекании газа через змеевик совершаемая над ним работа равна нулю. Не кажется ли вам странным этот результат? Легко придумать такой опыт, в котором над газом работа совершается, а теплообмена с окружающей средой нет. Действительно, попробуем при помощи компрессора прокачивать через змеевик газ в вакуум. Для того чтобы процесс можно было считать стационарным, сечение выходного отверстия сделаем много меньше сечения входного. Змеевик теплоизолируем от окружающей среды.

Совершаемая компрессором над газом работа положительна и равна полной совершаемой над газом работе, ибо, как уже отмечалось, выходя в вакуум, газ работы не совершает. Так как нет обмена теплотой, налицо противоречие с утверждением о том, что при адиабатическом протекании работа равна нулю.

Это противоречие возникло потому, что при прокачивании газа в вакуум происходят и такие энергетические превращения, которые были совершенно несущественны в разобранном выше примере. Действительно, первый закон термодинамики использовался в виде $Q + A = \Delta U$, где U — внутренняя энергия газа. Поэтому при использовании такой формулировки первого закона термодинамики заранее молчаливо предполагается, что в рассматриваемых процессах не происходит изменения механической энергии системы, т. е. не меняется потенциальная энергия газа как целого во внешнем поле, не меняется и кинетическая энергия движения газа как целого, не возникает в газе никаких макроскопических потоков. Теперь уже становится ясно, что при прокачивании газа в вакуум возникает макроскопический направленный поток, кинетическую энергию которого необходимо учитывать. Работа компрессора в этом случае как раз и определяет кинетическую энергию этого потока.

Если вход и выход змеевика расположены на разной высоте, то в уравнении закона сохранения энергии необходимо учитывать и изменение потенциальной энергии газа в поле тяжести, подобно тому как это делалось в гидродинамике при выводе уравнения Бернулли.

Измерение теплоемкости газа. Змеевик, помещенный в калориметр, можно использовать для измерения теплоемкости газа. Дело в том, что непосредственное измерение теплоемкости газа при постоянном объеме C_V затруднительно, так как для заключенного в сосуд газа масса, а следовательно, и теплоемкость всегда малы по сравнению с их значениями для сосуда и калориметра. Через змеевик можно пропустить большую массу газа, так что температура воды в

калориметре может заметно измениться. Таким образом удается преодолеть отмеченную выше трудность.

Зная теплоемкость калориметра с водой, по изменению его температуры можно подсчитать количество переданной газу теплоты Q (в пересчете на один моль прошедшего газа). Это количество теплоты связано с изменением температуры газа формулой (9), что позволяет рассчитать величину C_p — теплоемкость газа при постоянном давлении. Теплоемкость C_V можно затем найти с помощью уравнения Майера: $C_V = C_p - R$.

Задача

В вертикальном цилиндрическом сосуде под массивным поршнем находится 1 моль идеального газа. Приложим к поршню некоторую силу и сожмем газ, настолько, чтобы эта сила совершила заданную работу A. Затем поршень отпускаем и через некоторое время он устанавливается в новом положении равновесия. Определить, на сколько градусов температура газа в конечном состоянии отличается от начальной температуры, считая газ адиабатически изолированным.

Решение. До приложения внешней силы поршень находился в механическом равновесии: разность сил давления газа под поршнем и атмосферы над ним уравновешивала действующую на поршень силу тяжести. Приложив дополнительную силу, мы нарушаем это равновесие, и поршень перемещается вниз. При перемещении поршня все действующие на него внешние силы совершают положительную работу: кроме заданной в условии работы A приложенной силы это еще работа действующей на поршень сверху силы атмосферного давления и работа силы тяжести. Работа всех этих сил равна изменению внутренней энергии газа и кинетической энергии поршня.

При сжатии адиабатически изолированного газа его температура и давление возрастают. Поэтому после прекращения действия приложенной силы газ начинает расширяться, и поршень будет перемещаться вверх. При этом сила атмосферного давления и сила тяжести совершают отрицательную работу. В конце концов кинетическая энергия поршня обратится в нуль, и он установится в новом положении равновесия. Это новое положение равновесия поршня будет расположено выше исходного на некоторую величину Δh.

Полная работа, складывающаяся из работы A приложенной силы, работы $A_{\text{тяж}}$ силы тяжести и работы $A_{\text{атм}}$ силы атмосферного давления на всем перемещении поршня от начального положения, в соответствии со сказанным выше, равна изменению внутренней энергии газа ΔU:

$$A + A_{\text{тяж}} + A_{\text{атм}} = \Delta U. \tag{10}$$

Так как конечное положение поршня на Δh выше начального, то очевидно, что

$$A_{\text{тяж}} = -mg \cdot \Delta h, \quad A_{\text{атм}} = -p_0 S \cdot \Delta h, \tag{11}$$

где m — масса поршня, S — его площадь, p_0 — атмосферное давление. Изменение внутренней энергии ΔU находящегося под поршнем одного моля идеального газа определяется только изменением его температуры ΔT:

$$\Delta U = C_V \cdot \Delta T. \tag{12}$$

Подставляя (11) и (12) в (10), получаем

$$A - (mg + p_0 S)\,\Delta h = C_V \cdot \Delta T. \qquad (13)$$

Во втором слагаемом левой части (13) вынесем за скобки площадь поршня S. Тогда $S \cdot \Delta h$ равно изменению объема газа ΔV, а $mg/S + p_0$ — давление газа p, одинаковое в начальном и конечном состояниях. Поэтому (13) переписывается в виде

$$A - p \cdot \Delta V = C_V \cdot \Delta T. \qquad (14)$$

Применяя уравнение Менделеева—Клапейрона к начальному и конечному состояниям газа, имеем

$$p \cdot \Delta V = R \cdot \Delta T. \qquad (15)$$

Теперь соотношение (14) принимает вид

$$A = (C_V + R) \cdot \Delta T.$$

Так как $C_V + R = C_p$, то для изменения температуры ΔT имеем: $\Delta T = A/C_p$.

- Можно ли применять первый закон термодинамики к процессам, в которых изменяется химический состав или количество вещества в рассматриваемой системе?
- Всегда ли подведение теплоты к системе приводит к увеличению ее внутренней энергии?
- По показаниям какого прибора — термометра или барометра — можно судить о внутренней энергии всего воздуха, находящегося в вашей комнате?
- При прохождении газа через змеевик мы составляем уравнение энергетического баланса, в котором не учитываем кинетическую энергию потока газа, считая малой его скорость. При этом мы не заботимся о законе сохранения импульса. Медленность протекания газа можно обеспечить, например, с помощью пористой перегородки внутри змеевика. Поясните, почему такая перегородка, обеспечивая сохранение импульса, не влияет на энергетический баланс.
- Почему теплоемкость газа при постоянном объеме C_V трудно измерить на опыте непосредственно? Как эту трудность можно преодолеть на практике?
- В приведенном решении задачи утверждалось, что конечное положение поршня расположено выше начального. Докажите, что это действительно так.
- В условии задачи предполагалось, что приложенная сила перемещает поршень вниз. Изменится ли ответ, если считать, что она перемещает поршень вверх? Опишите процессы, происходящие в системе в этом случае.

§ 18. Второй закон термодинамики

Первый закон термодинамики — один из самых общих и фундаментальных законов природы. Не известно ни одного процесса, где хоть

в какой-то мере наблюдалось бы его нарушение. Если какой-либо процесс запрещен первым законом, то можно быть абсолютно уверенным в том, что он никогда не произойдет. Однако этот закон не дает никаких указаний о том, в каком направлении развиваются процессы, удовлетворяющие принципу сохранения энергии.

Поясним это примерами.

Направление тепловых процессов. Первый закон термодинамики ничего не говорит о том, в каком направлении происходит теплообмен между приведенными в тепловой контакт телами, находящимися при разных температурах. Как уже обсуждалось выше, теплообмен происходит так, что температуры выравниваются и вся система стремится к состоянию теплового равновесия. Но первый закон не был бы нарушен, если бы, наоборот, передача теплоты происходила от тела с низкой температурой к телу с более высокой при условии, что полный запас внутренней энергии оставался бы неизменным. Однако повседневный опыт показывает, что само собой это никогда не происходит.

Другой пример: при падении камня с некоторой высоты вся кинетическая энергия его поступательного движения исчезает при ударе о землю, но при этом увеличивается внутренняя энергия самого камня и окружающих его тел, так что закон сохранения энергии, разумеется, не оказывается нарушенным. Но первому закону термодинамики не противоречил бы и обратный процесс, при котором к лежащему на земле камню перешло бы от окружающих предметов некоторое количество теплоты, в результате чего камень поднялся бы на некоторую высоту. Однако никто никогда не наблюдал таких самопроизвольно подскакивающих камней.

Неравноценность разных видов энергии. Вдумываясь в эти и другие подобные примеры, мы приходим к выводу, что первый закон термодинамики не накладывает никаких ограничений на направление превращений энергии из одного вида в другой и на направление перехода теплоты между телами, требуя только сохранения полного запаса энергии в замкнутых системах. Между тем опыт показывает, что *разные виды энергии не равноценны* в отношении способности превращаться в другие виды.

Механическую энергию можно целиком превратить во внутреннюю энергию любого тела независимо от того, какова была его температура. Действительно, любое тело можно нагреть трением, увеличивая его внутреннюю энергию на величину, равную совершенной работе. Точно так же электрическая энергия может быть целиком превращена во внутреннюю, например при прохождении электрического тока через сопротивление.

Для обратных превращений внутренней энергии в другие виды существуют определенные ограничения, состоящие в том, что запас внутренней энергии ни при каких условиях не может превратиться

целиком в другие виды энергии. С отмеченными особенностями энергетических превращений связано направление протекания процессов в природе. Второй закон термодинамики, отражающий направленность естественных процессов и налагающий ограничения на возможные направления энергетических превращений в макроскопических системах, представляет собой, как и всякий фундаментальный закон, обобщение большого числа опытных фактов.

Чтобы яснее представить себе физическое содержание второго закона термодинамики, рассмотрим подробнее вопрос об обратимости тепловых процессов.

Обратимые и необратимые процессы. Если достаточно медленно изменять условия так, чтобы при этом скорость протекающего в рассматриваемой системе процесса была значительно меньше скорости релаксации, то такой процесс будет физически представлять собой цепочку близких друг к другу равновесных состояний. Поэтому такой процесс описывается теми же самыми макроскопическими параметрами, что и состояние равновесия. Эти медленные процессы называются *равновесными* или *квазистатическими*. При таких процессах систему можно характеризовать такими параметрами, как давление, температура и т. д. Реальные процессы являются неравновесными и могут считаться равновесными с большей или меньшей точностью.

Рассмотрим следующие примеры.

Пусть газ находится в цилиндрическом сосуде, закрытом поршнем. Если выдвигать поршень с конечной скоростью, то расширение газа будет необратимым процессом. Действительно, как только поршень будет выдвинут, давление газа непосредственно у поршня будет меньше, чем в других частях цилиндра. Такой процесс нельзя провести обратимо через те же промежуточные состояния, так как при вдвигании поршня обратно с конечной скоростью вблизи поршня будет происходить не разрежение газа, а его сжатие. Таким образом, быстрое расширение или сжатие газа дает пример необратимого процесса.

Чтобы расширить газ строго обратимым образом, нужно выдвигать поршень бесконечно медленно. При этом давление газа будет в каждый момент во всем объеме одинаковым, состояние газа будет зависеть от положения поршня, а не от направления его движения, и процесс будет обратимым.

Наиболее ярко необратимость процесса расширения газа проявляется тогда, когда расширение происходит в пустоту без совершения механической работы.

Необратимыми являются все процессы, сопровождающиеся теплообменом между телами, имеющими разные температуры. Необратимость такого теплообмена особенно отчетливо видна на примере выравнивания температур тел, приведенных в соприкосновение.

Необратимыми являются процессы, при которых механическая энергия переходит во внутреннюю при наличии трения, о чем часто говорят как о выделении теплоты благодаря трению. В отсутствие трения все механические процессы протекали бы обратимо.

Таким образом, равновесные обратимые процессы являются абстракцией, и на практике из-за существования трения и теплообмена не встречаются. Однако исследование равновесных процессов в термодинамике дает возможность указать, как следует проводить процессы в реальных системах, чтобы получить наилучшие результаты.

Различные формулировки второго закона термодинамики. Исторически открытие второго закона термодинамики было связано с изучением вопроса о максимальном коэффициенте полезного действия тепловых машин, проведенным французским ученым Сади Карно. Позднее Р. Клаузиус и У. Томсон (лорд Кельвин) предложили различные по виду, но эквивалентные формулировки второго закона термодинамики.

Согласно формулировке Клаузиуса, невозможен процесс, единственным результатом которого был бы переход теплоты от тела с более низкой температурой к телу с более высокой температурой.

Томсон сформулировал второй закон термодинамики следующим образом: невозможен периодический процесс, единственным конечным результатом которого было бы совершение работы за счет теплоты, взятой от одного какого-то тела.

Выражение «единственным результатом» в этих формулировках означает, что никаких других изменений, кроме указанных, ни в рассматриваемых системах, ни в окружающих их телах не происходит. Условная схема такого рода процесса, запрещенного постулатом Клаузиуса, показана на рис. 56, а процесса, запрещенного постулатом Томсона, — на рис. 57.

В формулировке Томсона второй закон термодинамики накладывает ограничения на превращение внутренней энергии в механическую. Из формулировки Томсона следует, что невозможно построить машину, которая совершала бы работу только лишь за счет получения теплоты из окружающей среды. Такая гипотетическая машина получила название вечного двигателя второго рода, так как вследствие неограниченности запасов внутренней энергии в земле, океане, атмосфере такая машина была бы для всех практических целей эквивалентна вечному двигателю.

Вечный двигатель второго рода не находится в противоречии с первым законом термодинамики, в отличие от вечного двигателя первого рода, т. е. устройства для совершения работы вообще без использования источника энергии.

Эквивалентность формулировок Клаузиуса и Томсона. Эквивалентность формулировок второго закона термодинамики, предло-

женных Клаузиусом и Томсоном, устанавливается простыми рассуждениями.

Предположим, что постулат Томсона несправедлив. Тогда можно осуществить такой процесс, единственным результатом которого было бы совершение работы за счет теплоты, взятой от единственного источника с температурой T. Эту работу можно было бы, например путем трения, снова целиком превратить в теплоту, передаваемую телу, температура которого выше, чем T. Единственным результатом такого составного процесса был бы переход теплоты от тела с температурой T к телу с более высокой температурой. Но это противоречило бы постулату Клаузиуса. Итак, постулат Клаузиуса не может быть справедливым, если неверен постулат Томсона.

Предположим теперь, что, наоборот, несправедлив постулат Клаузиуса, и покажем, что при этом постулат Томсона также не может вы-

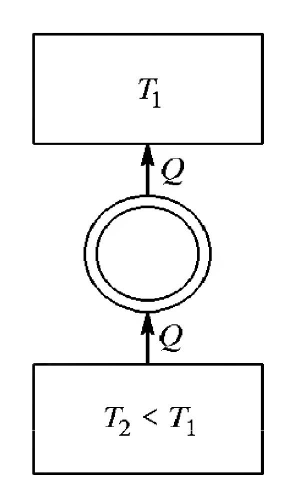


Рис. 56. Принципиальная схема гипотетического устройства, в котором нарушается постулат Клаузиуса

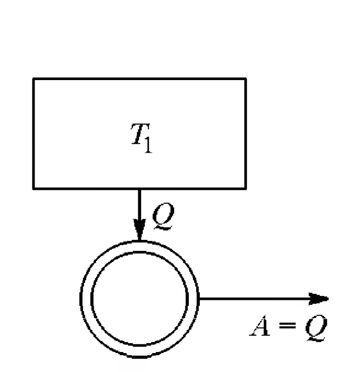


Рис. 57. Принципиальная схема гипотетического устройства, в котором нарушается постулат Томсона

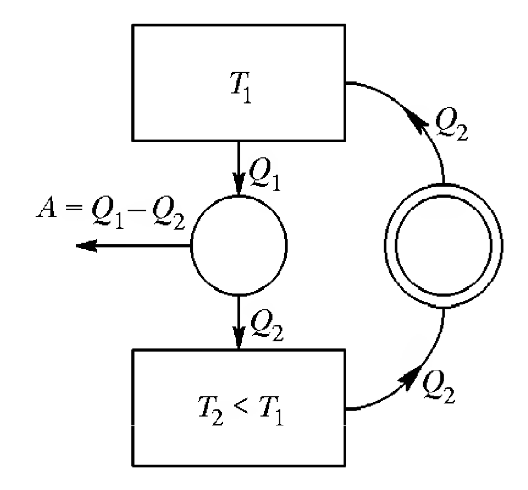


Рис. 58. Комбинируя с тепловой машиной устройство, изображенное на рис. 56, в котором нарушается постулат Клаузиуса, получаем систему, в которой на-рушается постулат Томсона

полняться. Построим обычную тепловую машину, которая будет работать, получая некоторое количество теплоты Q_1 от нагревателя, отдавая Q_2 холодильнику и превращая разность $Q_1 - Q_2$ в работу (рис. 58).

Поскольку постулат Клаузиуса предполагается неверным, можно осуществить процесс, единственным результатом которого будет переход количества теплоты, равного Q_2, от холодильника к нагревателю. Схематически это показано в правой части рис. 58. В результате нагреватель будет отдавать рабочему телу тепловой машины теплоту Q_1 и получать при процессе, противоречащем постулату Клаузиуса, теплоту Q_2, так что в целом он будет отдавать количество теплоты, равное $Q_1 - Q_2$. Именно такое количество теплоты машина превращает в работу. В холодильнике в целом никаких изменений вообще

не происходит, ибо он отдает и получает одно и то же количество теплоты Q_2. Теперь видно, что, комбинируя действие тепловой машины и процесс, противоречащий постулату Клаузиуса, можно получить процесс, противоречащий постулату Томсона.

Таким образом, постулаты Клаузиуса и Томсона либо оба верны, либо оба неверны, и в этом смысле они эквивалентны. Их справедливость для макроскопических систем подтверждается всеми имеющимися экспериментальными фактами.

Принцип Каратеодори. Физическое содержание второго закона термодинамики в формулировках Клаузиуса и Томсона выражается в виде утверждения о невозможности конкретных тепловых процессов. Но можно дать и такую формулировку, которая не конкретизирует вида процесса, невозможность которого утверждается этим законом. Такая формулировка называется принципом Каратеодори. Согласно этому принципу вблизи каждого равновесного состояния любой термодинамической системы существуют другие равновесные состояния, недостижимые из первого адиабатическим путем.

Покажем эквивалентность формулировки Томсона и принципа Каратеодори. Пусть произвольная термодинамическая система квазистатически переходит из некоторого состояния *1* в близкое состояние *2*, получая некоторое количество теплоты $Q > 0$ и совершая работу A'. Тогда в соответствии с первым законом термодинамики

$$Q = U_2 - U_1 + A'. \qquad (1)$$

Вернем систему адиабатически из состояния 2 в состояние 1. Тогда в таком обратном процессе теплообмен отсутствует, и первый закон термодинамики дает

$$0 = U_1 - U_2 + A'', \qquad (2)$$

где A'' — совершаемая системой работа. Складывая (1) и (2), получаем

$$Q = A' + A'' > 0. \qquad (3)$$

Соотношение (3) показывает, что в таком циклическом процессе система, возвратившись в исходное состояние, превратила в работу всю полученную теплоту. Но это невозможно, согласно второму закону термодинамики в формулировке Томсона. Значит, такой циклический процесс неосуществим. Первый его этап всегда возможен: на этом этапе к системе просто подводится теплота, и никаких других условий не накладывается. Поэтому невозможным здесь является только второй этап, когда по условию система должна возвращаться в исходное состояние адиабатически. Другими словами, состояние *1* адиабатически недостижимо из близкого к нему состояния *2*.

Принцип адиабатической недостижимости означает, что практически все реальные физические процессы происходят с теплообменом: адиабатические процессы — это редкое исключение. Рядом с каждым равновесным состоянием есть множество других, переход в которые обязательно требует теплообмена, и лишь в немногие из них можно попасть адиабатически.

На основе приведенных формулировок второго закона термодинамики можно получить результаты Карно для максимально возможного коэффициента полезного действия тепловых машин. Для тепловой машины, совершающей цикл между нагревателем с фиксированной температурой T_1 и холодильником с температурой T_2, коэффициент полезного действия не может превышать значения

$$\eta = \frac{T_1 - T_2}{T_1}. \qquad (4)$$

Наибольшее значение η, определяемое формулой (4), достигается у тепловой машины, совершающей обратимый цикл, независимо от того, что используется в качестве рабочего тела. Это утверждение, называемое обычно *теоремой Карно*, будет доказано ниже.

Цикл является обратимым, если он состоит из обратимых процессов, т. е. таких, которые можно провести в любом направлении через одну и ту же цепочку равновесных состояний.

Единственным обратимым циклическим процессом, который можно осуществить между нагревателем и холодильником с фиксированными температурами, является так называемый *цикл Карно*, состоящий из двух изотерм и двух адиабат. Для идеального газа такой цикл изображен на рис. 59. На участке *1—2* газ имеет температуру, равную температуре нагревателя T_1 и изотермически расширяется, получая теплоту Q_1 от нагревателя. При этом газ совершает положительную работу, равную полученной теплоте. На участке *2—3* газ расширяется адиабатически, и при этом его температура понижается от T_1 до значения, равного температуре холодильника T_2. Совершаемая газом на этом участке работа равна убыли его внутренней энергии. На следующем участке *3—4* газ изотермически сжимают. При этом он отдает холодильнику количество теплоты Q_2, равное совершаемой над ним при сжатии работе. На участке *4—1* газ адиабатически сжимают до тех пор, пока его температура не повысится до

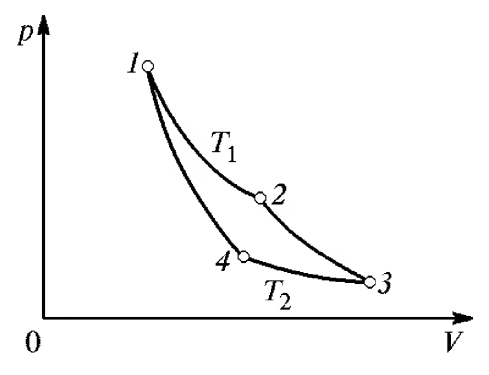


Рис. 59. Цикл Карно на p—V-диаграмме идеального газа

значения T_1. Увеличение внутренней энергии газа при этом равно работе внешних сил, совершаемой при сжатии газа.

Цикл Карно является единственным замкнутым процессом, который можно осуществить обратимым образом. В самом деле, адиабатические процессы обратимы, если их проводить достаточно медленно, т. е. квазистатически. Изотермические процессы — это единственные процессы с теплообменом, которые могут быть проведены обратимым образом. При любом другом процессе температура рабочего тела изменяется и, согласно второму закону термодинамики, теплообмен с нагревателем или холодильником не может быть обратимым: обмен теплотой при наличии конечной разности температур носит характер приближения к тепловому равновесию и не является равновесным процессом.

Разумеется, обмен теплотой в отсутствие разности температур происходит бесконечно медленно. Поэтому обратимый цикл Карно продолжается бесконечно долго и мощность тепловой машины при максимально возможном КПД, определяемом формулой (4), стремится к нулю. Процессы в любой реальной машине обязательно содержат необратимые звенья, и, следовательно, ее КПД всегда меньше теоретического предела (4).

Условия получения максимальной работы. Преобразование внутренней энергии в механическую, как следует из второго закона термодинамики, не может быть произведено полностью. Для того чтобы превратить в механическую энергию максимально возможную часть внутренней энергии, необходимо использовать исключительно обратимые процессы. Для иллюстрации рассмотрим следующий пример. Пусть имеется некоторое тело, не находящееся в состоянии теплового равновесия с окружающей средой, например идеальный газ в цилиндре с поршнем, имеющий температуру T_1, более высокую, чем температура окружающей среды T (рис. 60). Каким образом можно получить наибольшую работу при условии, что в конечном состоянии газ должен занимать тот же объем, что и в начальном?

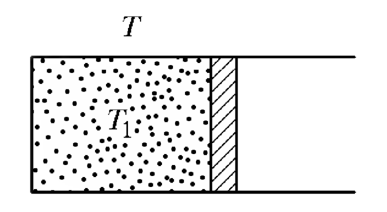


Рис. 60. К получению максимальной работы

Если бы температура газа была равна температуре окружающей среды, т. е. газ находился бы в тепловом равновесии с окружением, то никакой работы вообще получить было бы невозможно. Превращение внутренней энергии в механическую может происходить только в том случае, когда начальное состояние всей системы не является равновесным.

Но при неравновесном начальном состоянии переход системы в состояние равновесия не обязательно сопровождается превращением внутренней энергии в механическую. Если просто привести газ в тепловой контакт с окружающей средой, не давая ему расширяться,

то газ остынет и никакой работы при этом совершено не будет. Поэтому для возможности совершения работы необходимо предоставить газу возможность расширяться, имея в виду, что потом его придется сжать, так как по условию в конечном состоянии газ должен занимать тот же объем, что и в начальном.

Для получения максимальной работы переход из начального состояния в конечное должен быть произведен обратимо. А это можно сделать, только используя адиабатические и изотермические процессы. Итак, газ следует адиабатически расширять до тех пор, пока его температура не станет равна температуре окружающей среды T, а затем изотермически сжать при этой температуре до исходного объема (рис. 61). Совершаемая газом при адиабатическом расширении *1—2* работа, как видно из рисунка, больше той работы, которую придется совершить над газом при изотермическом сжатии *2—3*. Максимальная работа, которую можно получить при переходе газа из состояния *1* в состояние *3*, равна площади заштрихованного на рис. 61 криволинейного треугольника *1—2—3*.

Изученные закономерности действия обратимого теплового двигателя позволяют рассмотреть принципы функционирования холодильной машины и теплового насоса. В *холодильной машине* все процессы происходят в обратном (по сравнению с тепловым двигателем) направлении (рис. 62). За счет совершения механической работы A от резервуара с более низкой температурой T_2 отнимается некоторое количество теплоты Q_2. При этом резервуару с более высокой темпера-

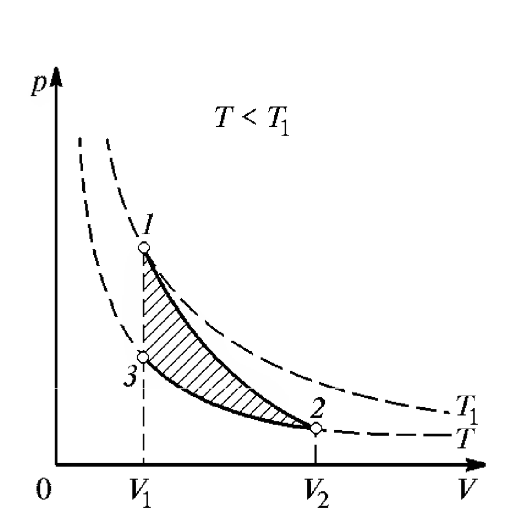


Рис. 61. Процесс получения максимальной работы на p—V-диаграмме

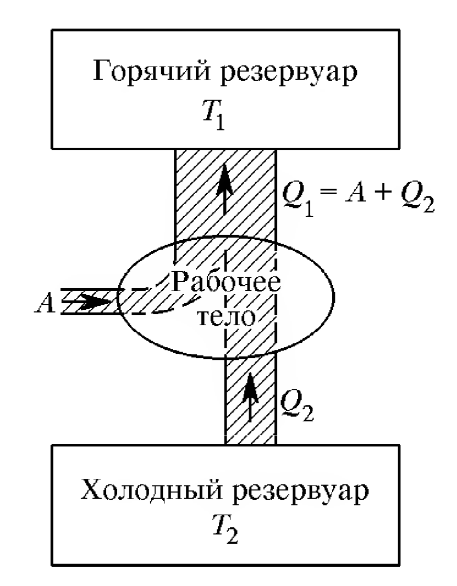


Рис. 62. Принципиальная схема холодильной машины

турой T_1, роль которого выполняет обычно окружающая среда, передается количество теплоты Q_1, равное сумме $A + Q_2$. Вследствие обратимости рассматриваемой машины для нее справедливо

соотношение

$$\frac{A}{Q_2+A}=\frac{T_1-T_2}{T_1}, \qquad (5)$$

которое в соответствии с (4) можно рассматривать как коэффициент полезного действия соответствующей тепловой машины.

Для холодильной машины наибольший интерес представляет количество теплоты Q_2, отнимаемое от охлаждаемого резервуара. Из (5) для Q_2 имеем

$$Q_2=\frac{A}{T_1/T_2-1}. \qquad (6)$$

График зависимости Q_2 от температуры окружающей среды T_1 (для обратимого процесса) изображен на рис. 63. Видно, что при $T_1 \gg T_2$ отнимаемая теплота $Q_2 \to 0$. Но при малой разности температур $T_1 - T_2$ отношение Q_2/A может принимать большие значения. Другими словами, эффективность холодильной машины при близких значениях T_1 и T_2 может быть весьма велика, так как количество теплоты Q_2, отнимаемое от охлаждаемых тел, может значительно превышать работу A, которую в реальных холодильных машинах совершает компрессор, приводимый в действие электродвигателем.

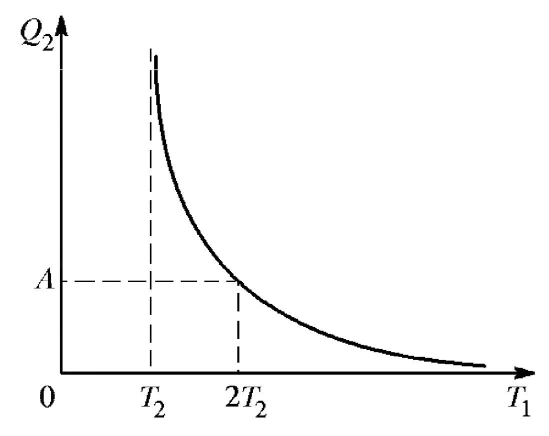


Рис. 63. Зависимость теплоты Q_2, отнимаемой у охлаждаемого резервуара, от температуры T_1 окружающей среды

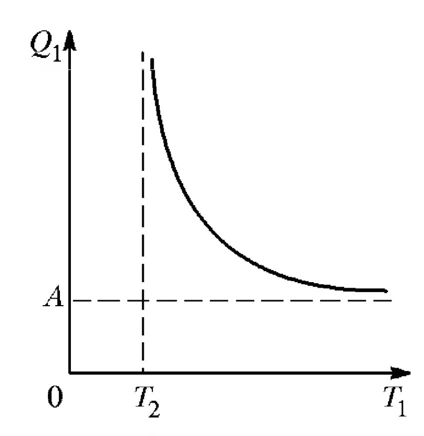


Рис. 64. Зависимость теплоты Q_1, поступающей в отапливаемое помещение, от его температуры T_1

В технической термодинамике для характеристики холодильной машины используется так называемый холодильный коэффициент ε, определяемый как отношение количества теплоты Q_2, взятого от охлаждаемых тел, к работе внешних сил A:

$$\varepsilon=\frac{Q_2}{A}=\frac{1}{T_1/T_2-1}. \qquad (7)$$

В отличие от КПД теплового двигателя (4), холодильный коэффициент ε может принимать значения, большие единицы. В реальных промышленных и бытовых установках $\varepsilon = 3$ и более. Как видно из (7), холодильный коэффициент тем больше, чем меньше различаются температуры окружающей среды и охлаждаемого тела.

Рассмотрим теперь работу теплового насоса, т. е. холодильной машины, работающей с целью нагревания горячего резервуара (отапливаемого помещения) за счет теплоты, отнятой от холодного резервуара (окружающей среды). Принципиальная схема теплового насоса идентична схеме холодильной машины (рис. 62). В отличие от холодильной машины для теплового насоса практический интерес представляет не Q_2, а Q_1 — количество теплоты, получаемое нагреваемым телом: $Q_1 = A + Q_2$. Для Q_1 аналогично (6) имеем

$$Q_1 = \frac{A}{1 - T_2/T_1}. \qquad (8)$$

График этой зависимости изображен на рис. 64. При $T_1 \gg T_2$ теплота $Q_1 \to A$. При малой разности $T_1 - T_2$ количество теплоты, передаваемое нагреваемому телу, может значительно превышать работу A. В этом случае становятся совершенно очевидными преимущества использования теплового насоса по сравнению с непосредственным отапливанием, когда работа A (например, электроэнергия) непосредственно превращается во внутреннюю энергию отапливаемого помещения.

В технической термодинамике для характеристики эффективности тепловых насосов вводится так называемый отопительный коэффициент $\varepsilon_{\text{отоп}}$, равный

$$\varepsilon_{\text{отоп}} = \frac{Q_1}{A} = \frac{1}{1 - T_2/T_1}. \qquad (9)$$

Приведенные формулы (7) и (9) справедливы для обратимых машин. Для реальных машин, где процессы полностью или частично необратимы, эти формулы дают оценку холодильного и отопительного коэффициентов.

Итак, при использовании теплового насоса отапливаемое помещение получает больше теплоты, чем при непосредственном отапливании. На это обстоятельство еще в 1852 г. обратил внимание У. Томсон, предложив идею так называемого динамического отопления, заключающуюся в следующем. Теплота, получаемая при сжигании топлива, используется не для непосредственного обогревания помещения, а направляется в тепловой двигатель для получения механической работы. С помощью этой работы приводится в действие тепловой насос, который и обогревает помещение. При малой разности температур окружающей среды и отапливаемого помещения последнее получает теплоты заметно больше, чем ее выделяется при сжигании топлива. Это может показаться парадоксальным.

В действительности никакого парадокса в тепловом насосе и динамическом отоплении нет, что становится совершенно ясным, если воспользоваться понятием качества внутренней энергии. Под качеством внутренней энергии понимается ее способность превращаться в другие виды. В этом смысле наивысшим качеством характеризуется энергия в механической или электромагнитной формах, так как ее можно полностью превратить во внутреннюю при любой температуре. Что касается внутренней энергии, то ее качество тем выше, чем выше температура тела, в котором она запасена. Всякий естественно идущий необратимый процесс, например переход теплоты к телу с более низкой температурой, ведет к обесцениванию внутренней энергии, к снижению ее качества. В обратимых процессах снижения качества энергии не происходит, поскольку все энергетические превращения могут идти в обратном направлении.

При обычном способе отапливания вся теплота, выделяющаяся при сжигании топлива при нагревании спирали электрическим током или получаемая от горячего резервуара и т. п., переходит в помещение в виде такого же количества теплоты, но при более низкой температуре, что представляет собой качественное обесценивание внутренней энергии. Тепловой насос или система динамического отопления устраняют непосредственный необратимый теплообмен между телами с разными температурами.

При работе теплового насоса или системы динамического отопления происходит повышение качества внутренней энергии, передаваемой отапливаемому помещению из окружающей среды. При малой разности температур, когда качество этой энергии существенно не увеличивается, ее количество становится больше, чем и объясняется высокая эффективность работы теплового насоса и динамического отопления в целом.

- Приведите примеры явлений, которые удовлетворяют закону сохранения энергии, но тем не менее никогда не наблюдаются в природе.
- В чем проявляется неравноценность разных видов энергии? Проиллюстрируйте эту неравноценность на примерах.
- Что такое обратимый тепловой процесс? Приведите примеры обратимых и необратимых процессов.
- Каким требованиям должна удовлетворять физическая система, чтобы механические процессы в ней протекали обратимо? Поясните, почему трение и диссипация механической энергии делают все процессы необратимыми.
- Приведите различные формулировки второго закона термодинамики. Докажите эквивалентность формулировок Клаузиуса и Томсона.
- Что означает принцип Каратеодори применительно к идеальному газу? Поясните ответ, используя $p-V$-диаграмму для изображения его состояния.

- Покажите, что физический смысл второго закона термодинамики заключается в установлении неразрывной связи между необратимостью реальных процессов в природе и теплообменом.
- Сформулируйте условия, при которых коэффициент полезного действия теплового двигателя, работающего по обратимому циклу, был бы близким к единице.
- Покажите, что цикл Карно — это единственный обратимый циклический процесс для двигателя, использующего два тепловых резервуара с фиксированными температурами.
- При обсуждении условий получения максимальной работы не учитывалось атмосферное давление, действующее на поршень снаружи. Как учет этого давления скажется на приведенных рассуждениях и на результате?
- Газ в цилиндре, закрытом поршнем, имеет такую же температуру, что и окружающий воздух, но более высокое (или более низкое) давление, чем давление в атмосфере. Какие процессы следует провести с газом, чтобы получить максимальную полезную работу за счет неравновесности системы? Изобразите эти процессы на $p-V$-диаграмме, считая газ в цилиндре идеальным.
- Газ в цилиндре, закрытом поршнем, имеет такое же давление, как и окружающий воздух, но более высокую (или более низкую) температуру. Какие процессы следует провести с газом, чтобы получить максимальную полезную работу за счет неравновесности системы? Изобразите их на $p-V$-диаграмме.
- Рассмотрите две различные схемы динамического отопления, в которых тепловая машина отдает теплоту Q_2 либо окружающей среде, либо отапливаемому помещению. Покажите, что в случае, когда все процессы обратимы, обе схемы имеют одинаковую эффективность. Какая схема окажется эффективнее в реальной системе, когда процессы нельзя считать полностью обратимыми?

§ 19. Методы термодинамики и их применения

В предыдущем параграфе была приведена без доказательства теорема Карно, выражающая максимально возможный коэффициент полезного действия теплового двигателя через температуры горячего и холодного тепловых резервуаров:

$$\eta = 1 - \frac{T_2}{T_1}. \tag{1}$$

Из второго закона термодинамики следует, что КПД тепловой машины, работающей по обратимому циклу, не зависит от того, какое вещество используется в качестве рабочего тела.

КПД всех обратимых тепловых машин одинаков. Убедиться в этом можно с помощью следующих рассуждений. Предположим

противное, т. е. что КПД η у одной из двух обратимых машин, работающих при заданных температурах T_1 и T_2, больше, чем у другой: $\eta > \eta'$. Допустим, что за один цикл обе машины отдают холодильнику одинаковые количества теплоты: $Q_2 = Q_2'$. Тогда машина, у которой КПД выше, за цикл превращает в работу большее количество теплоты: $Q_1 > Q_1'$ и $A_1 > A_1'$.

Пусть теперь машина с большим КПД работает в прямом, а с меньшим КПД — в обратном направлении, т. е. как холодильная машина. За счет совершения работы A' внешними телами от холодного резервуара отнимается теплота Q_2' и передается горячему резервуару теплота $Q_1' = Q_2' + A'$. В результате их совместного действия за цикл получается следующий результат. Холодильник получает от рабочего тела первой машины столько же теплоты, сколько отдает рабочему телу второй ($Q_2 = Q_2'$), т. е. в холодильнике никаких изменений не происходит. Поскольку $A_1 > A_1'$, то лишь части работы A_1 достаточно для приведения в действие холодильной машины. Нагреватель отдает теплоту Q_1, а получает меньшую теплоту Q_1', т. е. в целом отдает только $Q_1 - Q_1'$. Итак, теплота $Q_1 - Q_1'$, взятая от нагревателя, целиком превращается в полезную работу $A_1 - A_1'$, а холодильнику при этом никакой теплоты не передается. Это противоречит второму закону термодинамики в форме постулата Томсона. Значит, предположение, что $\eta > \eta'$, неверно.

Точно такие же рассуждения покажут, что для обратимых машин окажется неверным и предположение, что $\eta' > \eta$. Мы приходим к выводу, что все обратимые машины имеют одинаковый КПД.

Приведенное доказательство теоремы Карно основывалось на предположении об обратимости обеих тепловых машин, означающем, что любую из них можно было заставить действовать в обратном направлении. Если одна из машин, например с КПД η, необратима, то из приведенных рассуждений следует только, что $\eta < \eta'$: КПД любой необратимой тепловой машины меньше КПД обратимой машины, работающей при тех же температурах T_1 и T_2.

Независимость КПД обратимой тепловой машины от рода рабочего тела позволяет для получения формулы (1) выбрать простейшее рабочее тело — идеальный газ.

Для расчета КПД цикла Карно с идеальным газом нам понадобятся выражение для работы, совершаемой идеальным газом в изотермическом процессе, и уравнение адиабаты идеального газа.

Работа идеального газа в изотермическом процессе. Пусть один моль идеального газа изотермически расширяется от объема V_1 до

объема V_2. Совершаемая при этом работа выражается определенным интегралом:

$$A = \int_{V_1}^{V_2} p(V)\, dV. \tag{2}$$

Давление газа $p(V)$ можно выразить с помощью уравнения состояния: $p = RT/V$. Поскольку в изотермическом процессе температура постоянна, то выражение (2) принимает вид

$$A = RT \int_{V_1}^{V_2} \frac{dV}{V} = RT \ln \frac{V_2}{V_1}. \tag{3}$$

Вывод уравнения адиабаты. Уравнение адиабаты идеального газа можно получить, используя первый закон термодинамики и уравнение состояния. Так как в адиабатическом процессе теплообмен не происходит, то первый закон термодинамики записывается в виде

$$dU + p\, dV = 0. \tag{4}$$

Приращение внутренней энергии $dU = C_V dT$, поскольку внутренняя энергия идеального газа не зависит от объема. Выражение для dT можно получить из уравнения состояния $pV = RT$:

$$p\, dV + V\, dp = R\, dT. \tag{5}$$

С учетом приведенных выражений для dU и dT уравнение (4) переписывается в виде

$$C_V V\, dp + C_p p\, dV = 0, \tag{6}$$

где мы воспользовались тем, что $C_V + R = C_p$. Введем обозначение для отношения теплоемкостей C_p и C_V:

$$\gamma = \frac{C_p}{C_V}. \tag{7}$$

Разделив обе части (6) на произведение pV, получаем дифференциальное уравнение, выражающее зависимость давления идеального газа от объема в адиабатическом процессе:

$$\frac{dp}{p} + \gamma \frac{dV}{V} = 0. \tag{8}$$

Интегрируя (8), получаем

$$pV^{\gamma} = C, \tag{9}$$

где константа C сохраняет свое значение в течение адиабатического процесса. Это и есть уравнение адиабаты идеального газа.

Его можно записать и в других переменных, например V и T. Для этого можно просто подставить в (9) $p = RT/V$ из уравнения состояния:

$$TV^{\gamma-1} = C_1, \tag{10}$$

Аналогично уравнение адиабаты можно записать и в переменных p и T.

КПД цикла Карно. Теперь у нас есть все необходимое для получения формулы (1). Рассмотрим цикл Карно с идеальным газом, показанный на рис. 59. Газ получает теплоту Q_1 от нагревателя с температурой T_1 на изотермическом участке *1—2*, и отдает холодильнику с температурой T_2 теплоту Q_2 на изотермическом участке *3—4*. На адиабатических участках *2—3* и *4—1* теплообмена нет. Поскольку внутренняя энергия идеального газа при изотермическом процессе не меняется, то теплота Q_1 равна совершаемой газом работе при изотермическом расширении от V_1 до V_2. Поэтому в соответствии с (3) имеем

$$Q_1 = RT_1 \ln \frac{V_2}{V_1}. \tag{11}$$

Аналогично для теплоты Q_2 можно написать

$$Q_2 = RT_2 \ln \frac{V_3}{V_4}. \tag{12}$$

Теперь для КПД цикла Карно имеем

$$\eta = \frac{Q_1 - Q_2}{Q_1} = 1 - \frac{T_2}{T_1} \frac{\ln (V_3/V_4)}{\ln (V_2/V_1)}. \tag{13}$$

Легко видеть, что отношение логарифмов в (13) равно единице. В самом деле, с помощью (10) для адиабат *2—3* и *4—1* имеем

$$T_1 V_2^{\gamma-1} = T_2 V_3^{\gamma-1}, \qquad T_1 V_1^{\gamma-1} = T_2 V_4^{\gamma-1}. \tag{14}$$

Из этих равенств следует, что $V_2/V_1 = V_3/V_4$. Итак, для КПД цикла Карно с идеальным газом получаем формулу (1):

$$\eta = 1 - \frac{T_2}{T_1}.$$

В силу доказанной выше теоремы это выражение для КПД через температуры T_1 и T_2 справедливо для цикла Карно с любым рабочим телом.

Формулу (1) можно использовать для определения термодинамической температуры, не зависящей от свойств конкретных термометрических тел. Чтобы однозначно определить термодинамическую шкалу, необходимо задать значение температуры в некото-

рой реперной точке. В качестве такой точки выбирается тройная точка воды.

Неравенство Клаузиуса. Приведенные результаты позволяют дать второму закону термодинамики количественную формулировку в виде некоторого неравенства. Для любого кругового процесса между тепловыми резервуарами с температурами T_1 и T_2 ($T_1 > T_2$) справедливо соотношение, называемое неравенством Клаузиуса:

$$\frac{Q_1 - Q_2}{Q_1} \leqslant \frac{T_1 - T_2}{T_1}. \tag{15}$$

Здесь Q_1 — теплота, полученная от резервуара с температурой T_1, а Q_2 — теплота, отданная резервуару с температурой T_2. Знак равенства достигается в случае обратимого процесса, которым, как мы видели, может быть только цикл Карно.

Энтропия как функция состояния. Из (15) следует, что

$$\frac{Q_1}{T_1} \leqslant \frac{Q_2}{T_2}. \tag{16}$$

Если условиться считать теплоту, получаемую рабочим телом, положительной, а отдаваемую — отрицательной, то (16) можно переписать в виде

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} \leqslant 0, \tag{17}$$

где $Q_1 > 0$, а $Q_2 < 0$.

Неравенство (17) можно обобщить на случай любого кругового процесса, в котором рабочее тело обменивается теплотой с несколькими тепловыми резервуарами с различными температурами T_1, T_2, T_3, ...:

$$\sum_i \frac{Q_i}{T_i} \leqslant 0. \tag{18}$$

Если температура на протяжении кругового процесса изменяется непрерывно, то сумма в (18) обычным образом превращается в интеграл по замкнутому пути, так как рабочее тело, пройдя через ряд промежуточных состояний, возвращается в исходное состояние:

$$\oint \frac{\delta Q}{T} \leqslant 0, \tag{19}$$

где под δQ понимается получаемая (или отдаваемая) теплота на элементарном участке кругового процесса, настолько малом, чтобы температуру можно было считать постоянной.

В случае обратимого кругового процесса в (19) фигурирует знак равенства, т. е. интеграл по замкнутому контуру равен нулю:

$$\oint \frac{\delta Q}{T} = 0. \tag{20}$$

Это означает, что под интегралом стоит дифференциал некоторой функции состояния термодинамической системы: сумма ее приращений вдоль замкнутого пути обращается в нуль при возвращении системы в исходное состояние. Ситуация здесь такая же, как и при введении понятия потенциальной энергии в механике, где независимость работы от формы траектории (а следовательно, равенство нулю работы по любому замкнутому контуру) свидетельствовала о существовании функции состояния механической системы — потенциальной энергии. Итак,

$$\frac{\delta Q}{T} = dS, \tag{21}$$

где S — некоторая функция состояния термодинамической системы. Ее называют *энтропией*. Энтропия в каждом состоянии системы имеет вполне определенное значение, не зависящее от того, каким образом (обратимо или необратимо) система попала в данное состояние. При обратимом процессе изменение энтропии на элементарном участке, где температура T считается постоянной, равно отношению количества передаваемой теплоты δQ к абсолютной температуре. Формула (21) позволяет для каждой конкретной системы с известным уравнением состояния найти явный вид этой функции. При необратимом процессе $dS > \delta Q/T$.

Задача

Идеальный газ находится в одной половине теплоизолированного сосуда, разделенного перегородкой на две равные части. Определить изменение энтропии газа, если перегородка внезапно разрушается и газ заполняет весь сосуд.

Решение. Поскольку газ не получает теплоты (сосуд теплоизолирован) и не совершает работы (он расширяется в вакуум и никакого перемещения окружающих тел не происходит), его внутренняя энергия остается неизменной. Внутренняя энергия идеального газа пропорциональна его температуре и не зависит от объема, поэтому температура газа в результате описанного процесса не изменяется. Происходящий с газом процесс необратим. Однако энтропия конечного состояния газа, являясь функцией состояния, не зависит от того, каким путем попал газ в это конечное состояние. Поэтому можно рассмотреть обратимый изотермический процесс, при котором газ расширяется, перемещая перегородку, пока не займет весь объем сосуда.

Количество теплоты, получаемое идеальным газом при изотермическом процессе, равно совершаемой им работе:

$$Q = A' = \nu RT \ln \frac{V_2}{V_1} = \nu RT \ln 2,$$

так как $V_2/V_1 = 2$.

Вследствие постоянства температуры изменение энтропии системы записывается в виде

$$\Delta S = \frac{Q}{T} = \nu R \ln 2.$$

Объединенное уравнение первого и второго законов. Используя (21), можно записать равенство, выражающее и первый, и второй законы термодинамики. Для обратимых процессов из (21) следует, что

$$\delta Q = T\,dS. \tag{22}$$

Подставляя в уравнение первого закона термодинамики

$$dU = \delta Q + \delta A, \tag{23}$$

количество теплоты δQ, выраженное с помощью (22) через приращение энтропии dS, получаем

$$dU = T\,dS + \delta A. \tag{24}$$

Если система изотропна и совершает только механическую работу, то $\delta A = -p\,dV$ и (24) записывается в виде

$$dU = T\,dS - p\,dV. \tag{25}$$

Формула (25), объединяющая первый и второй законы термодинамики, называется фундаментальным равенством Гиббса.

Свободная энергия. Соотношение (25) выражает приращение внутренней энергии при изменении состояния системы. В качестве независимых переменных, задающих состояние термодинамической системы, могут использоваться разные макроскопические параметры. В (25) такими независимыми переменными выступают объем V и энтропия S. В практических приложениях использовать энтропию как независимую переменную неудобно. Более удобной физической величиной была бы температура, которую можно, как и объем, измерять непосредственно.

Перейти к независимым переменным T и V можно следующим образом. Введем еще одну функцию состояния F, связанную с внутренней энергией U соотношением

$$F = U - TS. \tag{26}$$

Функцию F называют *свободной энергией* системы. Для установления ее физического смысла запишем ее приращение при малом изменении состояния системы:

$$dF = dU - T\,dS - S\,dT. \tag{27}$$

Подставляя сюда dU из (25), получаем

$$dF = -S\,dT - p\,dV. \tag{28}$$

Видно, что независимыми переменными, определяющими состояние, для функции F являются температура и объем. Если изменить состояние системы изотермически, т. е. считать $dT = 0$, то, как видно из (28),

$$dF_{T=\text{const}} = -p\,dV.$$

Величина $-p\,dV$ представляет собой работу внешних сил над системой при изменении ее объема на dV. Таким образом, работа внешних сил при изотермическом процессе равна изменению свободной энергии системы.

Термодинамика диэлектриков и магнетиков. Полученные выше термодинамические соотношения позволяют исследовать энергетические превращения в самых разных физических системах. В частности, можно рассмотреть роль тепловых явлений в различных электромагнитных процессах. Оказывается, что, как и в случае реальных механических систем, практически любые энергетические превращения связаны с необратимым выделением теплоты.

Пусть, например, заполненный диэлектриком конденсатор заряжается от некоторого источника питания. При изменении заряда конденсатора на dq внешние силы (источник питания) совершают над ним работу

$$\delta A = \varphi\,dq, \tag{29}$$

где через φ обозначено напряжение на конденсаторе, т. е. разность потенциалов между его обкладками. Это напряжение в каждый момент пропорционально заряду конденсатора q:

$$\varphi = q/C = \frac{q}{\varepsilon C_0}. \tag{30}$$

В этом выражении $C = \varepsilon C_0$ — емкость конденсатора, равная произведению диэлектрической проницаемости ε диэлектрика в зазоре между обкладками на емкость C_0 того же конденсатора в отсутствие диэлектрика. В нашем случае основное соотношение (24), выражающее первый и второй законы термодинамики, записывается в виде

$$dU = T\,dS + \varphi\,dq, \tag{31}$$

или, после подстановки φ из (30),

$$dU = T\,dS + \frac{q\,dq}{\varepsilon C_0}. \tag{32}$$

Отметим, что внутренняя энергия U рассматриваемой системы (заряженного конденсатора) включает в себя и электростатическую энергию. В принципе с помощью (32) можно было бы получить выражение для этой внутренней энергии, сначала проинтегрировав

(32) при постоянной энтропии, и добавив затем изменение внутренней энергии при постоянном заряде. Однако практически сделать это затруднительно, поскольку при адиабатическом процессе ($S = \text{const}$) меняется температура системы, а вместе с ней и диэлектрическая проницаемость ε.

Обойти эту трудность можно, воспользовавшись вместо (32) аналогичным соотношением для свободной энергии F. В рассматриваемом случае вместо (28) имеем для dF:

$$dF = -S\,dT + \frac{q\,dq}{\varepsilon C_0}. \tag{33}$$

Это соотношение уже легко проинтегрировать при $T = \text{const}$, так как при постоянной температуре диэлектрическая проницаемость ε также неизменна. В результате получаем

$$F(T,\,q) = F(T) + \frac{q^2}{2\varepsilon C_0}. \tag{34}$$

Здесь $F(T)$ — свободная энергия системы в отсутствие электрического заряда. Мы не рассматриваем ее зависимость от объема, поскольку считаем, что в интересующих нас процессах изменение объема мало.

С помощью полученного выражения (34) можно вернуться к внутренней энергии. Из (26) имеем

$$U = F + TS, \tag{35}$$

где энтропию S можно найти с помощью (33):

$$S = -\frac{\partial F}{\partial T} = S(T) + \frac{q^2}{2\varepsilon^2 C_0}\frac{d\varepsilon}{dT}. \tag{36}$$

Здесь $S(T)$ — энтропия системы в отсутствие заряда (незаряженного конденсатора), получаемая в результате дифференцирования первого члена в правой части (34). Подставляя (34) и (36) в (35), получаем

$$U(T,\,q) = U(T) + \frac{q^2}{2\varepsilon C_0}\left(1 + \frac{T}{\varepsilon}\frac{d\varepsilon}{dT}\right). \tag{37}$$

Первое слагаемое в правой части (37) — это внутренняя энергия незаряженного конденсатора $U(T) = F(T) + TS(T)$, явное выражение для которой нас здесь не интересует. Второе слагаемое дает вклад во внутреннюю энергию системы, связанный с наличием заряда на конденсаторе.

В электростатике при выводе выражения для энергии заряженного конденсатора мы пренебрегали тепловыми процессами и считали, что энергия заряженного конденсатора равна работе электрических сил, совершаемой при его зарядке. При этом для энергии конденсатора получалось выражение $q^2/(2\varepsilon C_0)$. Из выражения (37)

видно, что это справедливо только в том случае, когда диэлектрическая проницаемость ε не зависит от температуры. В общем случае энергетические превращения в процессе зарядки конденсатора имеют более сложный характер.

Предположим, что процесс зарядки конденсатора происходит изотермически. В этом случае изменение внутренней энергии, определяемое вторым слагаемым в правой части (37), не равно работе внешних сил $q^2/(2\varepsilon C_0)$. Когда диэлектрическая проницаемость ε убывает с ростом температуры (т. е. $d\varepsilon/dT < 0$), запасаемая конденсатором зависящая от q энергия меньше работы внешних сил. Это означает, что в процессе адиабатической зарядки температура конденсатора будет повышаться, а при разрядке — понижаться. К выводу о понижении температуры мы пришли, рассматривая обратимые процессы. Поэтому такой эффект может наблюдаться лишь тогда, когда необратимым выделением джоулевой теплоты в веществе можно пренебречь.

Совершенно аналогично можно рассмотреть тепловые процессы, сопровождающие намагничивание и размагничивание магнетиков. Адиабатическое размагничивание парамагнитных солей используется в криогенной технике как метод получения наиболее низких температур, близких к абсолютному нулю.

Зависимость внутренней энергии от объема. Информацию о связи различных термодинамических характеристик можно получать из общих соотношений типа (25) и (28) с помощью формальных математических приемов. Например, с помощью этих соотношений легко найти закон зависимости внутренней энергии системы от ее объема. Действительно, разделив (25) на приращение объема dV, находим

$$\left(\frac{\partial U}{\partial V}\right)_T = T\left(\frac{\partial S}{\partial V}\right)_T - p. \qquad (38)$$

Но из (28) следует, что $(\partial S/\partial V)_T = (\partial p/\partial T)_V$. Поэтому (38) переписывается в виде

$$\left(\frac{\partial U}{\partial V}\right)_T = T\left(\frac{\partial p}{\partial T}\right)_V - p. \qquad (39)$$

Если известно уравнение состояния $p = p(V, T)$, то (39) позволяет найти явный вид зависимости внутренней энергии от объема. Например, для идеального газа $pV = RT$, $(\partial p/\partial T)_V = R/V = p/T$, и $(\partial U/\partial V)_T = 0$: внутренняя энергия идеального газа не зависит от объема.

- Проведите доказательство теоремы Карно с помощью таких рассуждений от противного, в которых использовался бы второй закон термодинамики в формулировке Клаузиуса.

- Почему при выводе уравнения адиабаты (9) величину γ в дифференциальном уравнении (8) можно считать постоянной?
- Почему при выборе общей формулы (1) для КПД теплового двигателя рассматривается частный случай — цикл Карно с идеальным газом?
- Поясните, почему неравенство Клаузиуса (15) или (16) можно рассматривать как альтернативную формулировку второго начала термодинамики.
- Поясните, каким образом из соотношения (20) следует вывод о существовании некоторой функции состояния.
- Почему выражение (21), связывающее приращение энтропии с передаваемой теплотой, справедливо только для обратимых процессов?
- Почему уравнение (24) или (25) можно рассматривать как соотношение, выражающее и первый и второй законы термодинамики?
- Покажите, что работа идеального газа при адиабатическом расширении может быть найдена по формуле

$$A = \frac{1}{\gamma - 1}(p_1 V_1 - p_2 V_2).$$

- При каких условиях работа внешних сил над термодинамической системой равна изменению свободной энергии?
- Почему при формулировке закона сохранения энергии в термодинамике используют понятие внутренней энергии, а не свободной энергии?
- Поясните, почему при адиабатической зарядке конденсатора с диэлектриком его температура повышается, когда $d\varepsilon/dT < 0$.
- Поясните, как методами термодинамики доказать, что внутренняя энергия идеального газа не зависит от его объема.

V. ОСНОВЫ МОЛЕКУЛЯРНО-КИНЕТИЧЕСКОЙ ТЕОРИИ

§ 20. Экспериментальные основания статистической механики

Основное положение молекулярно-кинетической теории строения вещества, вытекающее из опытных фактов, заключается в том, что атомы и молекулы, из которых состоят все макроскопические тела, находятся в состоянии непрерывного хаотического теплового движения.

Тепловое движение молекул. Наиболее убедительным опытным фактом, наглядно подтверждающим хаотический характер теплового движения и зависимость интенсивности этого движения от температуры, является *броуновское движение.*

Впервые это явление наблюдал английский ботаник Р. Броун в 1827 г., рассматривая в микроскоп взвешенные в воде крошечные сферические частички — споры плауна. Броуновское движение можно наблюдать и в газе. Его совершают, например, взвешенные в воздухе мелкие частицы пыли или дыма. Молекулярно-кинетическая теория броуновского движения была создана А. Эйнштейном только в 1905 г. В настоящее время термин «броуновское движение» используется в более широком смысле. Броуновским движением называют, в частности, дрожание стрелок чувствительных приборов, которое происходит из-за теплового движения молекул в самом приборе и в окружающей среде.

Наблюдая в микроскоп за движением взвешенных в жидкости мелких частиц, можно обнаружить, что каждая частица совершает хаотическое движение. Представление о характере блужданий частицы можно получить, если фиксировать ее положение в поле зрения измерительного микроскопа через равные промежутки времени. Соединив последовательные положения частицы прямыми, мы получим ломаную линию, подобную изображенной на рис. 65. Направления соседних участков ломаной линии составляют всевозможные углы друг с другом, так что не удается подметить никакой закономерности в изменении направления ломаной линии. Чем короче промежутки времени, через которые фиксируется положение частицы, тем более изломанной будет выглядеть «траектория» частицы: точки A, B, C, ... фиксируют положе-

ние частицы через 30 секунд, а точки *1*, *2*, *3*, ..., соединенные штриховыми линиями, фиксируют ее положение через каждые 5 секунд.

Наблюдение броуновского движения. Если наблюдать за движением сразу нескольких взвешенных в жидкости частиц, то можно заметить, что они движутся то в одну сторону, то в противоположные, то под углом друг к другу. Отсюда можно сделать вывод, что наблюдаемое броуновское движение не связано с перемещением потоков жидкости, так как в этом случае соседние частицы всегда перемещались бы вместе. На опыте же в условиях термодинамического равновесия никакой согласованности в движении соседних частиц не наблюдается, движутся они совершенно независимо друг от друга.

Изменяя температуру, при которой производится опыт, можно заметить, что с увеличением температуры интенсивность броуновского движения растет, с понижением температуры оно замирает.

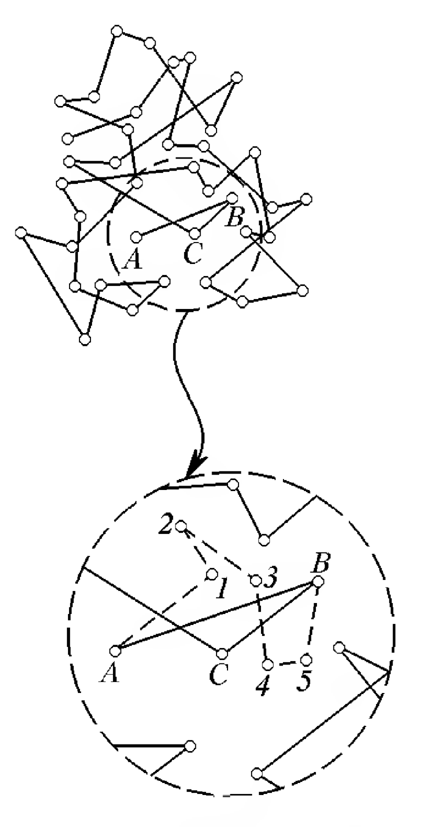


Рис. 65. Броуновское движение

Такой характер движения позволяет предположить, что броуновская частица движется под действием ударов, получаемых от молекул жидкости, в которой она находится. Если при этом считать, что тепловое движение молекул жидкости является хаотическим, то можно объяснить все наблюдаемые на опыте закономерности броуновского движения.

Закономерности броуновского движения. На первый взгляд могло бы показаться, что совершенно хаотический, беспорядочный характер ударов отдельных молекул должен был бы приводить к тому, что броуновская частица, масса которой во много раз больше массы молекулы, вообще не должна была бы заметно перемещаться. В самом деле, действие ударов, полученных броуновской частицей с одной стороны, должно полностью компенсироваться ударами с противоположной стороны. В такой ситуации, казалось бы, броуновская частица может только «дрожать» на месте. Ошибка такого рассуждения заключается в том, что случайный процесс подменяется, по существу, регулярным чередованием воздействий с противоположных сторон. Но такое чередование уже не является случайным процессом, а обладает высокой степенью упорядоченности. Степень упорядоченности такого чередования не отличается от степени упорядоченности процесса, в котором все испытываемые частицей толчки происходят в од-

ном направлении. Если, например, результат одного толчка характеризуется некоторым расстоянием l, то результат последовательности N упорядоченных толчков пропорционален величине Nl. Если же последовательность этих N толчков носит случайный характер, то их результат пропорционален $l\sqrt{N}$. Покажем это.

Будем с помощью измерительного микроскопа определять расстояние, на которое броуновская частица удаляется от начала координат за время t, многократно повторяя этот опыт. Всякий раз мы будем получать разные значения этого расстояния, однако в большинстве опытов будут получаться близкие друг к другу значения и лишь изредка заметно отличающиеся от остальных. Можно ввести среднее расстояние, на которое уходит частица от начала координат. Направления перемещений в отдельных опытах могут быть совершенно различными, поскольку все направления равновероятны.

Зависимость среднего перемещения от времени. Задача состоит в том, чтобы найти зависимость от времени среднего расстояния, которое будем обозначать символом $\langle R\rangle$.

Разделим интересующее нас время наблюдения t на большое число равных малых промежутков Δt таких, что в течение каждого промежутка частица испытывает огромное число ударов со стороны молекул жидкости. По существу, такое рассуждение означает многократное повторение опыта по измерению среднего расстояния, пройденного частицей за время Δt, причем каждый раз мы совмещаем начало координат с положением частицы в конце предыдущего промежутка времени Δt. Другими словами, это такой же опыт, как и рассмотренный выше, только осуществляемый за промежуток времени Δt, а не t. Поскольку и за промежуток Δt частица испытывает огромное число ударов, все приведенные выше рассуждения остаются в силе: направление перемещения за каждый «шаг» Δt совершенно произвольно и никак не связано с направлением перемещений в другие промежутки, а расстояние, проходимое частицей за Δt, будет примерно одинаковым для большинства промежутков.

Для простоты будем считать это расстояние одинаковым для всех шагов и обозначим его через L.

Пусть в результате N таких последовательных шагов частица оказалась в точке с радиусом-вектором $\mathbf{R}_N$. Тогда после очередного шага она попала в точку

$$\mathbf{R}_{N+1} = \mathbf{R}_N + \mathbf{L}_{N+1}, \tag{1}$$

где $\mathbf{L}_{N+1}$ — вектор перемещения за $(N+1)$-й шаг, имеющий произвольное направление и определенную длину L. Расстояние частицы от начала координат после $(N+1)$-го шага равно

$$R_{N+1} = \sqrt{R_N^2 + 2R_N L\cdot\cos\varphi + L^2}. \tag{2}$$

Здесь φ — угол между векторами $\mathbf{R}_N$ и $\mathbf{L}_{N+1}$. Найти среднее значение правой части этого выражения затруднительно, ибо усреднять нужно квадратный корень, а в общем случае среднее значение функции не равно этой функции от среднего значения аргумента: $\langle f(x)\rangle \neq f(\langle x\rangle)$. Легко заметить, что если возвести (1) или (2) в квадрат:

$$R_{N+1}^2 = R_N^2 + 2R_N L\cdot\cos\varphi + L^2, \tag{3}$$

то среднее значение квадрата смещения может быть легко найдено. Поэтому будем использовать для характеристики удаления броуновской частицы от начала координат не $\langle R\rangle$, а $\sqrt{\langle R^2\rangle}$. Усредняя левую и правую части (3) и учитывая, что угол φ с равной вероятностью принимает любые значения от 0 до 2π, т. е. $\langle\cos\varphi\rangle = 0$, получаем

$$\langle R_{N+1}^2\rangle = \langle R_N^2\rangle + L^2. \tag{4}$$

Используя метод математической индукции, на основе соотношения (4) легко показать, что

$$\langle R_N^2\rangle = NL^2. \tag{5}$$

Таким образом, среднее значение квадрата смещения пропорционально числу шагов, а поскольку шаги совершаются за одинаковые промежутки времени Δt, то

$$\langle R^2\rangle = \alpha t. \tag{6}$$

Это, конечно, не означает, что среднее смещение пропорционально времени. Броуновское движение частицы таково, что средний *квадрат* смещения растет пропорционально времени. Другими словами, квадратный корень из $\langle R^2\rangle$ растет со временем пропорционально $\sqrt{t}$. Эта величина, т. е. $\sqrt{\langle R^2\rangle}$, называемая средним квадратичным значением R, не равна среднему значению расстояния $\langle R\rangle$ частицы от начала координат спустя промежуток времени t, которое мы хотели определить. Однако можно показать, что эти величины отличаются только постоянным множителем. Поэтому среднее расстояние броуновской частицы от начала координат также пропорционально $\sqrt{t}$:

$$\langle R\rangle = \beta\sqrt{t}. \tag{7}$$

Совершенно очевидно, что коэффициенты α и β в формулах (6) и (7) зависят от интенсивности теплового движения молекул жидкости, удары которых приводят к броуновскому движению взвешенной частицы, т. е. в конечном счете от температуры.

Эксперимент и статистическая механика. Изучение броуновского движения сыграло большую роль в развитии молекулярно-кинетической теории строения вещества. Именно броуновское движение не только принесло неопровержимое доказательство реальности атомов и молекул, но и позволило впервые подсчитать количество молекул

в макроскопическом объеме вещества, т. е. определить значение постоянной Авогадро: $N_А = 6{,}02 \cdot 10^{23}$ моль$^{-1}$. Таким образом, было окончательно установлено, что тепловая форма движения материи обусловлена хаотическим движением атомов или молекул, из которых состоят макроскопические тела. Последнюю точку в этом вопросе поставили опыты французского физика Ж. Перрена, выполненные в начале XX века. Тем самым была подведена надежная экспериментальная база под статистическую механику, которая изучает свойства макроскопических систем, исходя из определенных модельных представлений о внутренней структуре вещества.

Постановка задачи в статистической механике. Задачей статистической механики является установление законов поведения макроскопических систем, состоящих из огромного числа частиц, на основе известных динамических законов поведения отдельных частиц. Другими словами, статистическая механика устанавливает связь между экспериментально измеряемыми макроскопическими величинами, характеризующими систему в целом, такими, как давление, объем, температура, напряженность электрического поля и т. д., и микроскопическими характеристиками системы, такими, как массы и заряды составляющих систему частиц, их координаты и импульсы и т. д.

Поясним сказанное на примере. Простейшей системой, состоящей из большого числа частиц, является газ, занимающий некоторый объем. С точки зрения механики состояние такой системы (т. е. ее микросостояние) определяется заданием положений и скоростей всех молекул газа, число которых в макроскопическом объеме огромно. Например, всего 1 см^3 воздуха при нормальных условиях содержит $2{,}7 \cdot 10^{19}$ молекул. Из-за движения молекул механическое состояние непрерывно изменяется. Однако опыт показывает, что при неизменных внешних условиях любая макроскопическая система рано или поздно приходит в стационарное состояние, при котором, несмотря на изменение механического состояния, такие макроскопические параметры, как, например, температура, плотность, давление, характеризующие макросостояние системы, остаются неизменными. Для изолированной макроскопической системы это будет состояние теплового равновесия.

Таким образом, определение состояния системы в статистической механике является гораздо менее детализированным, чем в механике, так как опирается лишь на небольшое число макроскопических параметров, измеряемых на опыте. В большинстве случаев такое *сокращенное описание* системы является вполне достаточным, ибо нас, как правило, совершенно не интересует детальная информация о движении отдельных молекул.

Но значения макроскопических параметров, разумеется, зависят от движения молекул, и задача статистической механики — выразить свойства системы в целом через характеристики отдельных мо-

лекул, т. е. перекинуть мост между макро- и микроскопическими описаниями системы. При этом требуется установить связь макроскопических параметров системы со средними значениями микроскопических величин и дать способ вычисления этих средних значений на основе законов движения отдельных молекул.

Статистическая механика и термодинамика. Напомним, что в отличие от молекулярно-кинетической теории термодинамический подход не опирается ни на какие модельные представления об атомно-молекулярной структуре вещества. Основные понятия термодинамики вводятся на основе физического эксперимента, и поэтому она оперирует только макроскопическими величинами: давлением, температурой, объемом и т. п. Термодинамический подход отличается большой общностью и простотой. Он дает возможность решать многие конкретные задачи, не требуя никаких сведений о свойствах атомов или молекул.

Недостатком термодинамического метода можно считать то, что при его использовании остается невыявленной связь между наблюдаемым явлением и обусловливающим это явление поведением молекул. Если мы, например, термодинамическим методом установим, что металлический стержень при нагревании должен удлиняться, а растянутый резиновый жгут — сокращаться, то мы не сможем объяснить, какие особенности строения вещества приводят к такому различию в поведении при нагревании. Если же это нас не удовлетворяет и мы хотим понять, почему так происходит, то мы должны обратиться к статистической механике, так как в рамках термодинамики невозможно вскрыть глубокий физический смысл макроскопических параметров и их связь с микроскопическими параметрами.

Статистическая механика и термодинамика долгое время развивались независимо, ибо термодинамика основывалась на экспериментальных фактах, в то время как в основе статистической механики лежали гипотезы об атомно-молекулярном строении вещества и кинетической природе теплоты, достоверность которых вызывала сомнение до тех пор, пока эти гипотезы не были подтверждены экспериментально. С тех пор отпала необходимость в резком разграничении между термодинамикой и молекулярно-кинетической теорией, и в настоящее время они фактически слились в единую науку — статистическую термодинамику.

Действительно, как мы увидим, наиболее полные представления о свойствах систем большого числа частиц дает совместное использование термодинамики и статистической механики.

- Объясните, почему броуновская частица не «топчется» на месте, а систематически смещается из начального положения, несмотря на то, что молекулы окружающей среды в равной мере бомбардируют ее со всех сторон.
- Почему среднее значение $\cos\varphi$ равно нулю, если аргумент φ принимает с равной вероятностью любые значения от 0 до 2π?

- Докажите, используя метод математической индукции, что средний квадрат смещения броуновской частицы за N шагов пропорционален числу шагов N.
- Как зависит среднее смещение броуновской частицы от числа шагов N? от времени наблюдения?
- Почему коэффициенты пропорциональности α и β в формулах (6) и (7) зависят от температуры?
- В чем состоит принципиальное отличие подхода к изучению макроскопических систем в термодинамике и в статистической механике?
- Как в статистической механике макроскопические параметры системы связаны с микроскопическими характеристиками используемых физических моделей строения вещества?

§ 21. Молекулярно-кинетическая теория идеального газа

Статистическая механика дает возможность установить связь между макроскопическими параметрами большой системы и средними значениями микроскопических величин, характеризующих отдельные молекулы. Проиллюстрируем это на простом, но важном примере уравнения состояния идеального газа.

В разреженном газе расстояния между молекулами во много раз превышают их размеры. Так как молекулы электрически нейтральны, их взаимодействие между собой проявляется только на малых расстояниях. Поэтому большую часть времени молекулы газа проводят в свободном движении, и их кинетическая энергия в среднем преобладает над потенциальной.

Физическая модель газа. Вместо того чтобы рассматривать реальный газ, введем его простейшую физическую модель — идеальный газ. Будем считать, что идеальный газ — это совокупность огромного числа одинаковых молекул, размеры которых пренебрежимо малы. В рамках такой модели считается, что молекулы движутся по законам классической механики и взаимодействуют между собой только во время столкновений, которые имеют характер упругого удара.

Вывод основного уравнения кинетической теории идеального газа. Газ заключен в сосуд, и в состоянии теплового равновесия никаких макроскопических движений в нем не происходит. Рассчитаем давление, оказываемое идеальным газом на стенку сосуда. Для этого нужно сопоставить этот макроскопический параметр с результатом бомбардировки стенки отдельными молекулами. Суммарное действие молекул на поверхность можно заменить одной непрерывно действующей силой, так как молекул очень много и их столкновения со стенкой происходят очень часто. Согласно законам динамики Ньютона, эта сила равна модулю изменения перпендикулярной к стенке составляющей импульса всех молекул, которые испытывают столкновения с ней за единицу времени.

Столкновение молекулы со стенкой. Хотя место и время удара каждой молекулы о стенку совершенно несущественны, начинать приходится с рассмотрения удара отдельной молекулы. Пока для простоты предположим, что молекулы сталкиваются со стенкой абсолютно упруго. Когда молекула отскакивает от стенки, проекция ее скорости на направление нормали к стенке меняет знак. Направим ось x по нормали к стенке (рис. 66). Обозначим через v_x проекцию скорости молекулы до удара, а через $v'_x = -v_x$ — после удара. Изменение импульса молекулы при столкновении со стенкой $m(v'_x - v_x)$ равно $-2mv_x$, а передаваемый стенке импульс равен $2mv_x$. Это относится к единичному столкновению.

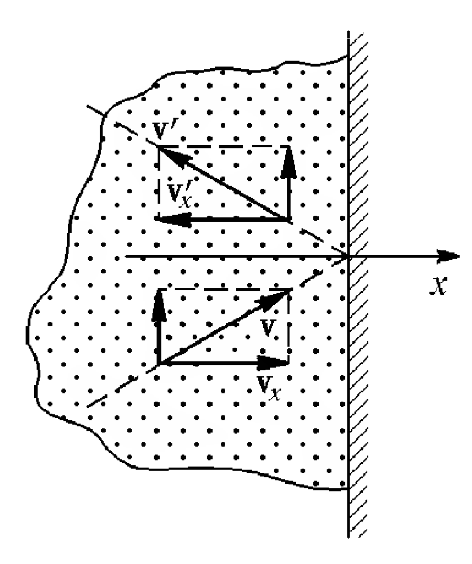


Рис. 66. Упругое столкновение молекулы газа со стенкой сосуда

О роли столкновений молекул. При подсчете давления газа можно не думать о столкновениях молекул между собой благодаря хаотичности их теплового движения: если в результате столкновения рассматриваемая молекула изменит свою скорость, то в большом ансамбле молекул в состоянии равновесия на смену ей непременно придет другая молекула, получившая такую же скорость в результате другого столкновения.

Так как давление газа зависит от его температуры и объема и не зависит от формы сосуда, возьмем для простоты сосуд в форме прямоугольного параллелепипеда. Двигаясь как свободная, молекула проходит расстояние l между противоположными стенками за время l/v_x, так что второе столкновение с той же стенкой произойдет спустя время $2l/v_x$. Следовательно, за достаточно большое время t молекула нанесет по данной стенке $v_x t/2l$ ударов, передавая ей каждый раз импульс $2mv_x$. В результате за время t эта молекула передаст стенке импульс, равный

$$2mv_x \cdot v_x t/2l.$$

Суммирование по всем молекулам. Для нахождения импульса, передаваемого стенке всеми N молекулами газа, нужно сложить результаты действия всех молекул. Силу давления на стенку мы получим, разделив полный передаваемый импульс на время t, а давление p — разделив эту силу на площадь S стенки:

$$p = \frac{m}{V} \sum_{i=1}^{N} v_{ix}^2. \qquad (1)$$

Здесь $V = Sl$ — объем сосуда, в котором заключен газ, а v_{ix} — проекция скорости i-й молекулы на перпендикулярную к стенке сосуда ось x. Обратим внимание на то, что сумма квадратов v_{ix}^2, деленная на число молекул N, представляет собой среднее значение квадрата проекции скорости одной молекулы:

$$\langle v_x^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} v_{ix}^2. \tag{2}$$

Поэтому формулу (1) можно переписать в виде

$$p = \frac{mN}{V} \langle v_x^2 \rangle. \tag{3}$$

Среднее значение квадрата скорости. Входящую в (3) величину $\langle v_x^2 \rangle$ можно выразить через среднее значение квадрата скорости молекулы $\langle v^2 \rangle$. В самом деле, $v^2 = v_x^2 + v_y^2 + v_z^2$. Для средних значений, очевидно, справедливо $\langle v^2 \rangle = \langle v_x^2 \rangle + \langle v_y^2 \rangle + \langle v_z^2 \rangle$. Так как в состоянии равновесия все направления в газе равноправны, то

$$\langle v_x^2 \rangle = \langle v_y^2 \rangle = \langle v_z^2 \rangle = \frac{1}{3} \langle v^2 \rangle. \tag{4}$$

Теперь выражение (3) принимает вид

$$p = \frac{1}{3} \frac{mN}{V} \langle v^2 \rangle, \tag{5}$$

или

$$p = \frac{1}{3} mn \langle v^2 \rangle, \tag{6}$$

где $n = N/V$ — среднее число молекул газа в единице объема, или концентрация молекул.

Выражение (5) или (6) называют *основным уравнением кинетической теории идеального газа.* Макроскопический параметр, характеризующий газ в целом, — давление p — выражен здесь через среднее значение микроскопического параметра — квадрата скорости отдельной молекулы.

Идеальный газ в статистической механике и термодинамике. Рассмотрим подробнее полученное уравнение. В его правой части можно увидеть среднюю кинетическую энергию одной молекулы

$$\langle \varepsilon \rangle = \frac{1}{2} m \langle v^2 \rangle. \tag{7}$$

Так как внутренняя энергия идеального газа состоит только из кинетической энергии хаотического теплового движения его молекул, то произведение $N\langle \varepsilon \rangle$ и представляет собой внутреннюю энергию U идеального газа из N молекул. Поэтому уравнение (5) можно переписать в виде

$$pV = \frac{2}{3} U. \tag{8}$$

Таким образом, в рассматриваемой статистической модели идеального газа произведение давления на объем равно (с точностью до множителя 2/3) внутренней энергии U.

В термодинамике идеальным газом был назван газ, для которого точно выполняется эмпирическое уравнение Клапейрона—Менделеева. Сравнивая (8) с уравнением Клапейрона—Менделеева

$$pV = RT, \tag{9}$$

получаем выражение для внутренней энергии одного моля идеального газа:

$$U = \frac{3}{2} RT. \tag{10}$$

Мы видим, что внутренняя энергия идеального газа пропорциональна термодинамической температуре T и не зависит от объема. Это значит, что рассмотренная здесь статистическая модель адекватна представлению об идеальном газе, введенному в термодинамике. Эта модель соответствует одноатомному газу, поскольку его молекулы рассматривались как материальные точки, вся кинетическая энергия которых сводится только к энергии поступательного движения. Для двухатомных и многоатомных молекул необходимо было бы учитывать еще и кинетическую энергию их вращения как целого и энергию колебаний входящих в них атомов.

Используя соотношение (10) для внутренней энергии одноатомного газа, можно получить явное выражение для его молярной теплоемкости C_V при постоянном объеме:

$$C_V = \frac{dU}{dT} = \frac{3}{2} R. \tag{11}$$

Видно, что это универсальная постоянная величина, не зависящая от химического состава газа, его молярной массы и других характеристик газа в условиях, когда к нему применима модель одноатомного идеального газа.

Физический смысл температуры. Соотношению (10) можно придать другой вид, если рассматривать не внутреннюю энергию U одного моля газа, а среднюю кинетическую энергию, приходящуюся на одну молекулу $\langle \varepsilon \rangle = U/N$:

$$\langle \varepsilon \rangle = \frac{3}{2} kT, \tag{12}$$

где универсальная постоянная

$$k = \frac{R}{N_A}, \tag{13}$$

равная отношению универсальной газовой постоянной к постоянной Авогадро N_A, называется *постоянной Больцмана* в честь выдающегося австрийского физика Людвига Больцмана, основоположника теории кинетических явлений.

Формула (2) позволяет установить физический смысл термодинамической температуры T — физической величины, которая была введена ранее феноменологически. Этот макроскопический параметр, как видно из (12), характеризует среднее значение кинетической энергии хаотического теплового движения одной молекулы в состоянии термодинамического равновесия.

Интересно отметить, что средняя энергия теплового движения молекул зависит только от температуры газа. При данной температуре средняя кинетическая энергия поступательного хаотического движения молекул не зависит ни от химического состава газа, ни от массы молекул, ни от давления газа, ни от объема, занимаемого газом.

Равнораспределение энергии по степеням свободы. Молекула одноатомного идеального газа, рассматриваемая как материальная точка, имеет три степени свободы. Равенство средних значений квадратов проекций скорости молекулы, выражаемое соотношением (4), позволяет сделать вывод о том, что на каждую степень свободы в состоянии термодинамического равновесия в среднем приходится одна и та же энергия. Эта энергия, как следует из (12), равна $kT/2$.

Оказывается, что этот результат имеет универсальный характер: средняя *кинетическая* энергия, приходящаяся на каждую степень свободы, одинакова и равна $kT/2$. Это утверждение относится не только к газам, оно справедливо для теплового движения молекул в жидкостях и твердых телах, ионов и электронов в плазме и даже для макроскопических тел, совершающих броуновское движение в результате хаотических ударов молекул окружающей среды.

Давление газа и температура. Соотношения (7) и (12) позволяют переписать основное уравнение кинетической теории идеального газа (6) так, что давление газа выражается только через его концентрацию и температуру:

$$p = nkT. \tag{14}$$

Смеси различных газов. Поскольку в состоянии теплового равновесия средняя кинетическая энергия молекул зависит только от температуры, то в смеси газов средние кинетические энергии молекул разных сортов одинаковы:

$$\frac{m_1\langle v_1^2\rangle}{2} = \frac{m_2\langle v_2^2\rangle}{2} = \frac{3}{2}\,kT. \tag{15}$$

Поэтому легкие молекулы движутся в среднем быстрее тяжелых. Полное давление смеси идеальных газов равно сумме давлений, которые имел бы каждый из газов, составляющих смесь, если удалить из сосуда остальные газы. В этом можно убедиться, буквально повторяя приведенный вывод уравнения (6) и учитывая, что импульс, передаваемый стенке молекулами каждого сорта, обусловливает то

давление, которое создавал бы один этот газ. Этот закон был открыт опытным путем Дальтоном и носит его имя.

- Как вы понимаете утверждение, что в разреженном газе из электрически нейтральных молекул их кинетическая энергия в среднем преобладает над потенциальной?
- Опишите модель идеального газа, используемую в молекулярно-кинетической теории. Какую роль играют в этой модели собственные размеры молекул и их столкновения между собой?
- Чем определяется импульс, передаваемый стенке молекулой в единичном соударении с ней?
- Сформулируйте основные этапы вывода уравнения (5) кинетической теории газа. Какие гипотезы положены в основу вывода?
- Приведите аргументы, подтверждающие заключение о равенстве средних значений квадратов проекций скоростей молекулы на оси x, y, z.
- Приведите соображения, свидетельствующие о соответствии рассмотренной статистической модели идеального газа и модели идеального газа, принятой в термодинамике.
- Поясните, почему формулы (8), (10) и (11) применимы только для одноатомного газа.
- Как средняя энергия хаотического теплового движения молекул связана с температурой системы?
- Каким образом в статистической механике раскрывается физический смысл температуры?
- Какой вывод о скоростях теплового движения молекул разной массы в смеси газов можно сделать на основе представлений о равнораспределении кинетической энергии по степеням свободы?

Δ **О микроскопических моделях.** Вернемся еще раз к приведенному выше выводу основного уравнения кинетической теории идеального газа. Прежде всего отметим то обстоятельство, что в механике и электродинамике при использовании различных физических моделей мы рассматривали характеристики изучаемой системы, доступные непосредственному наблюдению и измерению в эксперименте. Но в статистической модели газа микроскопические параметры, характеризующие отдельные молекулы, непосредственно не наблюдаются, а с опытом сравниваются только их усредненные по всему ансамблю молекул значения, через которые выражаются макроскопические параметры. Поэтому особое значение здесь приобретает обоснованность предположений относительно микроскопической картины рассматриваемых явлений.

Рассмотрим с этих позиций допущения, которые были сделаны при выводе уравнения (5) или (6).

Другой вывод основного уравнения. Прежде всего могут показаться не вполне убедительными соображения, приведенные для

обоснования возможности пренебречь столкновениями молекул. Поэтому приведем несколько иной вывод, в котором этот вопрос исследуется более тщательно. Будем, как и раньше, удар молекулы о стенку считать абсолютно упругим, так что передаваемый стенке при единичном ударе импульс равен $2mv_x$. Учтем теперь, что ударяющиеся о стенку молекулы имеют разные значения проекций скорости v_x. Рассмотрим те молекулы, у которых проекция скорости на ось x лежит в малом интервале значений от v_x до $v_x + \Delta v_x$. Пусть число таких молекул в единице объема равно $\Delta n(v_x)$. За время Δt до стенки долетят и столкнутся с ней только те из них, которые находятся внутри слоя толщиной $v_x \Delta t$, прилегающего к участку стенки площади S (рис. 67).

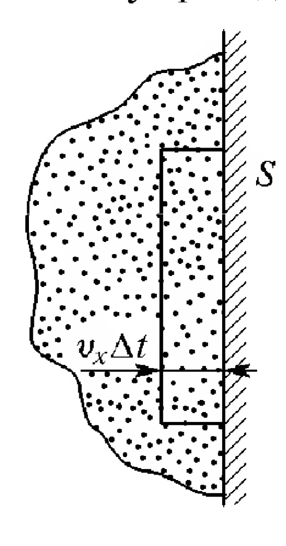


Рис. 67. К вычислению давления газа

Промежуток времени Δt можно выбрать настолько малым, чтобы толщина слоя $v_x \Delta t$ была много меньше длины свободного пробега молекул. Тогда столкновений молекул между собой в этом слое практически не будет. Итак, число ударов, наносимых рассматриваемыми молекулами за время Δt, равно $\Delta n(v_x) v_x S \Delta t$, а передаваемый при этом стенке импульс равен

$$2mv_x \Delta n(v_x) v_x S \Delta t.$$

Отсюда давление на стенку Δp, создаваемое этой группой молекул, равно

$$\Delta p = 2mv_x^2 \Delta n(v_x). \qquad (16)$$

Полное давление, создаваемое всеми молекулами, получим, просуммировав (16) по всем группам молекул, скорости которых направлены к стенке, т. е. по всевозможным значениям $v_x > 0$:

$$p = 2m \sum v_x^2 \Delta n(v_x) \quad (v_x > 0). \qquad (17)$$

Вследствие хаотичности теплового движения в состоянии равновесия число летящих к стенке молекул со скоростью, лежащей в интервале от v_x до $v_x + \Delta v_x$, в среднем равно числу летящих от стенки молекул со скоростью от $-v_x$ до $-(v_x + \Delta v_x)$, т. е. $\Delta n(v_x) = \Delta n(-v_x)$.

Так как под знаком суммы в (17) стоит квадрат проекции скорости, то сумма только по положительным значениям v_x равна половине суммы по всевозможным v_x:

$$p = m \sum v_x^2 \Delta n(v_x). \qquad (18)$$

Легко сообразить, что сумма в (18) связана со средним значением квадрата проекции скорости молекулы на ось x. В самом деле, среднее значение v_x^2 по совокупности из n молекул определяется формулой (2). Но в (18) фактически стоит та же самая сумма, только суммирование производится не по отдельным молекулам, а по группам молекул в единице объема, имеющих почти одинаковые значения v_x. Поэтому (18) можно переписать в виде

$$p = mn\langle v_x^2\rangle, \tag{19}$$

где n есть среднее число молекул в единице объема с любыми скоростями, т. е. концентрация. Дальнейшие рассуждения не отличаются от использованных ранее. В силу равноправия всех направлений при хаотическом тепловом движении $\langle v_x^2\rangle = \langle v^2\rangle/3$, и, следовательно, формула (19) принимает вид

$$p = \frac{1}{3}\, mn\langle v^2\rangle,$$

что совпадает с формулой (6).

Еще раз о роли столкновений. Теперь мы можем более четко сформулировать соображения о том, что столкновения молекул между собой не влияют на полученный результат. При подсчете числа ударов о площадку мы считали, что все молекулы из выделенной группы со скоростями от v_x до $v_x + \Delta v_x$, лежащие в слое толщиной $v_x\Delta t$, достигают стенки в течение промежутка времени Δt. Но на самом деле некоторые из них не долетят до стенки, так как в результате столкновений могут изменить направление скорости. Но кроме таких столкновений, в результате которых молекулы «уходят» из выделенной группы, в газе происходят столкновения, пополняющие эту группу молекул. В состоянии теплового равновесия, когда в газе отсутствуют макроскопические потоки, среднее число молекул в каждой группе не меняется со временем, несмотря на частые столкновения молекул. Поэтому столкновения не могут изменить среднего результата ударов молекул о стенки сосуда, выражаемого формулой (6).

Еще один вывод основного уравнения. В обоих приведенных выводах основного уравнения кинетической теории идеального газа фигурировала проекция скорости на направление нормали к стенке сосуда, и давление газа выражалось через среднее значение квадрата этой проекции $\langle v_x^2\rangle$ (формулы (3) и (19)). Затем значение $\langle v_x^2\rangle$ выражалось через $\langle v^2\rangle$ благодаря эквивалентности всех направлений в изотропном газе. Можно построить рассуждения таким образом, чтобы с самого начала учесть эту изотропность.

Поскольку давление газа не зависит от формы сосуда, возьмем сосуд сферической формы, в котором сначала имеется только одна молекула. На рис. 68 показано сечение этого сосуда плоскостью, проходящей через траекторию молекулы и центр сосуда. При абсолютно упругом столкновении со стенкой изменение импульса молекулы, как видно из этого рисунка, дается выражением $2mv\cdot\cos\alpha$. Именно такой импульс передается стенке по нормали к ней при каждом столкновении, так как угол α при последовательных упругих столкновениях со стенкой остается прежним. Так как давление в газе определяется модулем импульса, передаваемого молекулами стенке по нормали к ней, то можно просуммировать эти модули для последовательных столкновений, невзирая на то, что при каждом соударении нормаль к стенке имеет свое направление. Как видно из рис. 68, расстояние, проходимое молекулой между любыми двумя последовательными соударениями со стенкой, равно $2R\cdot\cos\alpha$, где R — радиус сосуда. Поэтому за промежуток времени t молекула нанесет по стенке $vt/(2R\cdot\cos\alpha)$ ударов и передаст по нормали к ней импульс

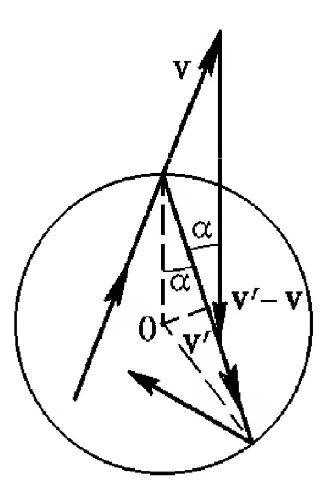


Рис. 68. Столкновение молекулы газа со стенкой сферического сосуда

$$2mv\cdot\cos\alpha\,\frac{vt}{2R\cdot\cos\alpha}=\frac{mv^2t}{R}.$$

Видно, что передаваемый по нормали импульс за время t не зависит от направления движения молекулы.

Для подсчета импульса, передаваемого стенке по нормали всеми молекулами газа, нужно сложить результаты действия всех молекул. Когда этих молекул много и они движутся хаотически, любой участок стенки испытывает одинаковое воздействие, и для нахождения давления p нужно просто разделить суммарный передаваемый по нормали импульс на время t и площадь поверхности сосуда $4\pi R^2$:

$$p=\frac{m}{4\pi R^3}\sum_{i=1}^{N}v_i^2. \tag{20}$$

Здесь сумму квадратов скорости всех молекул можно заменить произведением полного числа молекул N на среднее значение квадрата их скорости $\langle v^2\rangle$:

$$p=\frac{mN\langle v^2\rangle}{4\pi R^3}. \tag{21}$$

Учитывая, что объем сферического сосуда равен $(4/3)\pi R^3$, видим, что (21) совпадает с полученной ранее формулой (5). Отме-

тим, что в этом выводе коэффициент 1/3 в выражении для p появляется не в результате усреднения квадрата проекции скорости v_x^2, а как прямое следствие трехмерности физического пространства.

О характере столкновения со стенкой. Во всех приведенных выводах основного уравнения предполагалось, что столкновение молекулы со стенкой сосуда происходит по законам упругого удара.

На самом деле это предположение также является несущественным. Чтобы убедиться в этом, рассмотрим газ в прямоугольном сосуде (рис. 69), в котором стенки A и B обладают разными свойствами: удар молекул газа о стенку A абсолютно упругий, а о стенку B — неупругий. Если бы давление газа на эти стенки было различным, то в отсутствие трения о подставку сосуд с газом пришел бы в движение под действием внутренних сил. Так как сосуд в движение не приходит, мы должны заключить, что в состоянии теплового равновесия давление газа не зависит от характера взаимодействия его молекул со стенками. Объясняется это тем, что в стационарном состоянии молекулы не накапливаются на стенках: сколько молекул прилипает к стенке при неупругом ударе, столько же от нее и улетает, причем в среднем с такой же скоростью, так как температуры стенки и газа в состоянии теплового равновесия одинаковы.

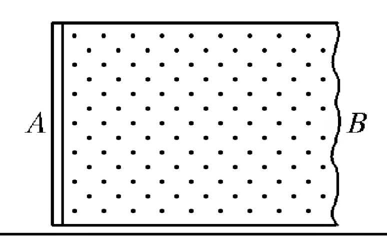


Рис. 69. Стенки сосуда A и B разные: удар молекул о стенку A абсолютно упругий, о стенку B — неупругий

В отличие от предположений о характере столкновений молекул между собой и со стенкой сосуда, принципиально важным оказывается допущение о том, что газ идеальный, т. е. его молекулы настолько малы, что их полный собственный объем мал по сравнению с объемом занимаемого газом сосуда, и что молекулы взаимодействуют только во время столкновений. Только для газа с такими свойствами и справедливо уравнение (6). Для реальных газов оно выполняется лишь приближенно. ▲

- В чем специфика молекулярных моделей, используемых в статистической механике, в сравнении с физическими моделями в механике макроскопических тел и электродинамике?
- Поясните, почему при выводе основного уравнения кинетической теории газа учет столкновений между молекулами не влияет на окончательный результат.
- Объясните качественно, почему в сферическом сосуде импульс, передаваемый молекулой по нормали к стенке за большое время, не зависит от угла падения на стенку. Рассмотрите также предельный случай скользящего падения, соответствующий движению молекулы по окружности радиуса R.

- Рассмотрите неупругое столкновение молекулы со стенкой сосуда. Какие предположения о механизме взаимодействия со стенкой необходимы в этом случае, чтобы модель описывала состояние теплового равновесия?
- В некоторых физических теориях рассматривается модель двумерного идеального газа, молекулы которого могут двигаться только в одной плоскости. Выведите основное уравнение кинетической теории такого газа, понимая под его давлением силу, действующую на единицу длины границы.

§ 22. Статистические распределения

Полный хаос, которым характеризуется тепловое движение молекул, имеет свои законы.

Законы хаоса. Несмотря на то, что каждая молекула газа при столкновениях с другими молекулами и со стенками сосуда все время изменяет свою скорость, макроскопическое состояние газа в термодинамическом равновесии не изменяется. Это позволяет считать, что в газе существует некоторое в среднем неизменное во времени *распределение молекул по скоростям*. Действительно, как мы видели, при данной температуре среднее значение квадрата скорости молекул имеет определенное значение. Однако среди молекул в данный момент времени есть и быстрые, и медленные, и можно поставить вопрос: сколько в среднем в газе молекул имеет то или иное значение скорости? Другой представляющий интерес вопрос — как найти среднюю скорость, не зная значений скорости отдельных молекул?

Определенное распределение молекул газа по скоростям устанавливается всегда, когда газ приходит в равновесие, независимо от того, каково было начальное состояние системы. Если даже в откачанный сосуд впустить струю газа, в которой все молекулы имеют почти одинаковые по модулю и направлению скорости, то спустя некоторое время в результате столкновений молекул направленное движение в газе перейдет в хаотическое, при котором все направления скоростей будут встречаться одинаково часто, а в распределении молекул по модулю скорости будет наблюдаться определенная закономерность.

Равновесное состояние газа характеризуется не только распределением молекул по скоростям, но и по координатам. В отсутствие внешних полей это распределение будет однородным, т. е. газ равномерно распределяется по всему объему сосуда: в любых равных макроскопических объемах внутри сосуда в среднем находится одинаковое число молекул.

А как обстоит дело при наличии действующего на молекулы поля, например поля тяжести? Хорошо известно, что давление воздуха убывает с высотой. Следовательно, убывает и концентрация молекул воздуха. Например, на высоте Эльбруса (5600 м) давление со-

ставляет лишь половину давления на уровне моря, т. е. концентрация молекул там уже вдвое меньше. Отсюда, конечно, не следует делать вывод, что на вдвое большей высоте совсем нет молекул воздуха, — самолеты летают и гораздо выше.

Распределение молекул по высоте. Найти закон распределения молекул газа с высотой в однородном поле тяжести можно из условия механического равновесия. Рассмотрим вертикальный столб газа с площадью основания S (рис. 70) и выделим в нем мысленно на высоте z слой толщиной Δz настолько малой, чтобы плотность газа ρ можно было считать в пределах этого слоя постоянной, но в то же время эта толщина должна быть такой, чтобы внутри выделенного слоя было много молекул и можно было бы говорить о производимом ими давлении.

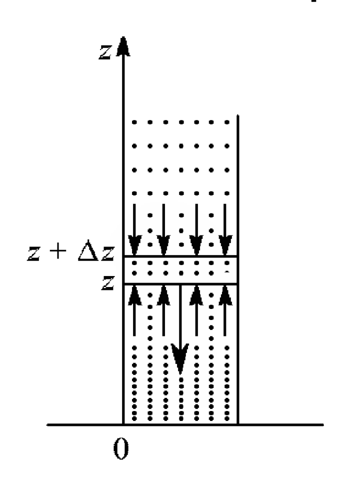


Рис. 70. Равновесие мысленно выделенного объема газа в поле тяжести

Применим к этому выделенному слою газа условие механического равновесия подобно тому, как это делалось в гидростатике для слоя жидкости, где мы, используя понятие давления, совершенно не интересовались его молекулярно-кинетической природой. Мы можем так поступать, ибо давление газа на стенку сосуда, рассматриваемое как результат передачи молекулами импульса стенке при столкновениях, и гидростатическое давление в газе или жидкости на опыте измеряются одинаково, одними и теми же приборами и, следовательно, представляют собой один и тот же макроскопический параметр рассматриваемой системы.

Условие механического равновесия выделенного слоя газа состоит в том, что действующая на него сила тяжести уравновешивается силами давления на верхнее и нижнее основания. В проекции на ось z (рис. 70) это условие записывается в виде

$$p(z)S - p(z + \Delta z)S - \rho g S \Delta z = 0.$$

Так как давление на высоте $z + \Delta z$ можно записать в виде

$$p(z + \Delta z) = p(z) + \Delta p,$$

то условие равновесия принимает вид

$$\Delta p = -\rho g \Delta z. \qquad (1)$$

Входящая в (1) плотность газа ρ зависит от давления. Выразим ее из уравнения Менделеева—Клапейрона для произвольной массы газа $m_г$

$$pV = \frac{m_г}{M} RT, \qquad (2)$$

где M — молярная масса. С помощью (2) получаем

$$\rho = \frac{m_\text{г}}{V} = \frac{Mp}{RT}. \tag{3}$$

Подставляем это выражение в (1) и переходим к пределу при $\Delta z \to 0$. Так как предел отношения $\Delta p / \Delta z$ при $\Delta z \to 0$ есть производная dp/dz, то получаем следующее дифференциальное уравнение для функции $p(z)$:

$$\frac{dp}{dz} = -\frac{Mg}{RT}\, p. \tag{4}$$

Это уравнение говорит о том, что производная искомой функции пропорциональна самой функции. Как известно, единственной функцией, обладающей таким свойством, является экспонента, и, следовательно, решение такого уравнения при постоянных g и T имеет вид

$$p(z) = C \exp\left(-\frac{Mgz}{RT}\right). \tag{5}$$

Значение постоянной C определяется из условия, что давление на высоте $z = 0$ равно величине p_0:

$$p(z) = p_0 \exp\left(-\frac{Mgz}{RT}\right). \tag{6}$$

Барометрическая формула. Формулу (6) можно переписать в несколько ином виде, учитывая, что молярная масса M равна произведению массы молекулы m на постоянную Авогадро $N_\text{А}$:

$$p(z) = p_0 \exp\left(-\frac{mgz}{kT}\right). \tag{7}$$

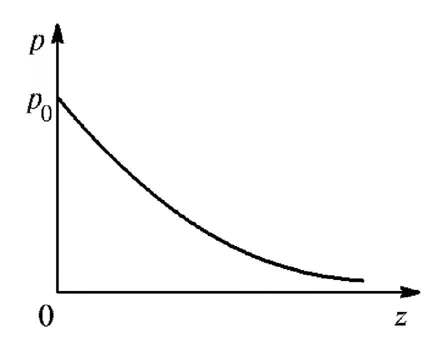


Рис. 71. Зависимость давления газа от высоты

Соотношение (6) или (7) называется *барометрической формулой*. Выражаемая ею зависимость давления газа p от высоты z графически представлена на рис. 71.

Отметим, что применимость барометрической формулы к реальной земной атмосфере весьма ограничена, так как атмосфера практически никогда не находится в состоянии теплового равновесия и ее температура меняется с высотой. Учитывая связь между давлением газа и концентрацией молекул

$$p = nkT$$

из (7) получаем распределение молекул по высоте во внешнем поле:

$$n(z) = n_0 \exp\left(-\frac{mgz}{kT}\right). \tag{8}$$

Распределение Больцмана. Легко заметить, что в числителе показателя экспоненты в (8) стоит потенциальная энергия молекулы, находящейся в поле тяжести на высоте z: $\varepsilon_\text{п}(z) = mgz$, т. е. эту формулу можно переписать в виде

$$n(z) = n_0 \exp\left(-\frac{\varepsilon_\text{п}(z)}{kT}\right). \qquad (9)$$

Этот полученный на конкретном примере результат имеет весьма общий характер. Формула (9) дает равновесное распределение молекул в пространстве в любом потенциальном поле и называется *распределением Больцмана.*

Общая теория равновесных статистических распределений была создана Гиббсом. Он показал, что в состоянии теплового равновесия при температуре T закон распределения молекул по любой характеризующей их состояние величине (координате, скорости, энергии) имеет экспоненциальный характер, причем в показателе экспоненты, как и в (9), стоит взятое со знаком минус отношение характерной энергии молекулы к величине kT, которая пропорциональна средней кинетической энергии хаотического движения молекул.

Распределение по проекции скорости. В частности, для распределения молекул газа по проекции скорости на какое-либо направление в показателе экспоненты стоит отношение зависящей от этой проекции части энергии молекулы к kT:

$$f(v_x) = a \exp\left(-\frac{\varepsilon(v_x)}{kT}\right) = a \exp\left(-\frac{mv_x^2}{2kT}\right). \qquad (10)$$

Величина $f(v_x)$ называется *функцией распределения* молекул по проекции скорости на ось x. Произведение $f(v_x)\Delta v_x$ равно среднему числу молекул газа в единице объема, у которых проекция скорости на ось x лежит в интервале от v_x до $v_x + \Delta v_x$, подобно тому как произведение $n(z)$ из формулы (9) на Δz дает среднее число молекул, z-координата которых лежит между z и $z + \Delta z$.

Функция распределения. Остановимся подробнее на смысле функции распределения $f(v_x)$. Прежде всего подчеркнем, что бессмысленно задавать вопрос о том, сколько молекул имеют строго определенное значение v_x, например ровно 500 м/с. Скорее всего, в данный момент во всем сосуде с газом не окажется ни одной молекулы с таким значением v_x, так как число молекул газа хоть и очень велико, но все же конечно, в то время как допустимых значений v_x бесконечно много. Поэтому имеет смысл говорить только о среднем числе молекул в единице объема, значение проекции скорости которых на ось x лежит в интервале от v_x до $v_x + \Delta v_x$, например от 500 до 501 м/с.

Нормировка функции распределения. Наглядное представление о законе распределения молекул по проекции скорости дает график функции $f(v_x)$, определяемой формулой (10), который приведен на рис. 72. Площадь заштрихованной полоски на этом рисунке, равная $f(v_x)\Delta v_x$, дает, как мы видели, среднее число молекул газа в единице объема, у которых проекция скорости v_x лежит в указанном интервале Δv_x. Теперь легко сообразить, что полная площадь, ограниченная графиком функции $f(v_x)$ и осью v_x, дает число молекул, у которых v_x имеет любые значения от $-\infty$ до ∞, т. е. дает полное число молекул в единице объема. Именно из этого условия, называемого условием нормировки функции распределения, и определяется постоянная a в формуле (10). Для ее нахождения нужно проинтегрировать функцию $f(v_x)$ по v_x от $-\infty$ до ∞ и приравнять результат концентрации молекул n. Это дает следующее значение постоянной a:

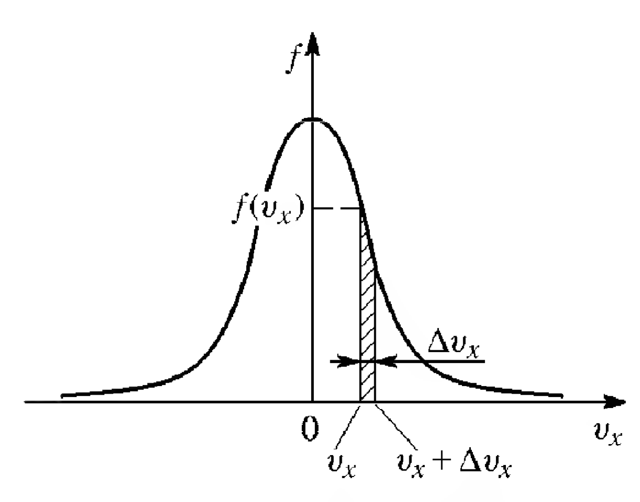


Рис. 72. График функции распределения $f(v_x)$

$$a = n\sqrt{\frac{m}{2\pi kT}}. \tag{11}$$

Функции распределения можно дать и несколько иную интерпретацию. Вместо того чтобы говорить о том, сколько в среднем молекул в единице объема имеют значение v_x в заданном интервале, можно говорить о том, какова вероятность того, что наугад выбранная молекула имеет значение v_x в этом интервале. Эту вероятность можно понимать как отношение среднего числа таких молекул к полному числу молекул в единице объема. Обозначая ее через $q(v_x)\Delta v_x$, можем написать

$$q(v_x) = \frac{1}{n} f(v_x) = \sqrt{\frac{m}{2\pi kT}} \exp\left(-\frac{mv_x^2}{2kT}\right). \tag{12}$$

Распределение по трем проекциям скорости. Из-за полной хаотичности теплового движения в состоянии равновесия можно считать, что все направления равноправны. В противном случае в газе существовало бы некоторое преимущественное направление движения молекул и, следовательно, существовал бы направленный поток газа, чего нет в состоянии теплового равновесия. Поэтому, если выбрать любые три взаимно перпендикулярных направления x, y, z, то

функции распределения молекул по проекциям скорости на эти направления будут иметь один и тот же вид (12).

Можно поставить вопрос, какова вероятность того, что наугад выбранная молекула газа будет иметь значения трех проекций ее скорости v_x, v_y и v_z в заданных интервалах: проекция на ось x в интервале от v_x до $v_x + \Delta v_x$, на ось y — от v_y до $v_y + \Delta v_y$, на ось z — от v_z до $v_z + \Delta v_z$. Вероятность иметь, например, v_x в заданном интервале не зависит от того, каковы значения двух других проекций скорости молекулы. Это значит, что вероятность того, что все три проекции скорости лежат в заданных интервалах, равна произведению вероятностей для каждой из проекций в отдельности:

$$q(v_x,v_y,v_z)\Delta v_x \Delta v_y \Delta v_z = q(v_x)\Delta v_x \cdot q(v_y)\Delta v_y \cdot q(v_z)\Delta v_z.$$

Теперь с помощью (12) можем написать

$$q(v_x,v_y,v_z) = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left[-\frac{m}{2kT}\left(v_x^2 + v_y^2 + v_z^2\right)\right]. \qquad (13)$$

В показателе экспоненты в (13) фактически стоит квадрат скорости молекулы, и показатель равен отношению ее кинетической энергии к kT. Таким образом, эта функция распределения зависит только от модуля скорости, но не от ее направления. Распределение по скоростям, как и следовало ожидать, оказывается равномерным по всем направлениям, т. е. изотропным в пространстве.

Распределение по модулю скорости. Теперь легко получить выражение для вероятности того, что наугад выбранная молекула имеет модуль скорости в заданном интервале от v до $v + \Delta v$, независимо от того, как эта скорость направлена. Для этого нужно найти, чему соответствует произведение $\Delta v_x \Delta v_y \Delta v_z$ для всех молекул, модуль скорости которых лежит в заданном интервале от v до $v + \Delta v$. Если построить систему координат v_x, v_y, v_z (рис. 73), то легко видеть, что молекулам с одинаковым значением модуля скорости v соответствуют точки на поверхности сферы радиуса v с центром в начале координат. Молекулам со скоростями в интервале от v до $v + \Delta v$ соответствует шаровой слой толщины Δv. Поэтому произведению $\Delta v_x \Delta v_y \Delta v_z$ в рассматриваемом случае соответствует объем этого шарового слоя, равный произведению площади поверхности сферы на толщину слоя, т. е. $4\pi v^2 \Delta v$. Выражение для вероятности того, что молекула имеет модуль скорости в

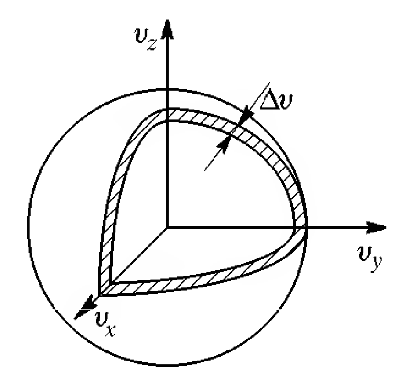


Рис. 73. Шаровой слой в пространстве скоростей

заданном интервале, равно произведению (13) на $4\pi v^2 \Delta v$:

$$q(v)\,\Delta v = 4\pi\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(-\frac{mv^2}{2kT}\right) v^2 \Delta v. \tag{14}$$

Если нас интересует среднее число молекул в единице объема с такими значениями модуля скорости, то мы должны умножить (14) на концентрацию газа n:

$$f(v)\,\Delta v = 4\pi n\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(-\frac{mv^2}{2kT}\right) v^2 \Delta v. \tag{15}$$

Зависимость функции $f(v)$ от скорости показана на рис. 74. Эта функция имеет максимум при значении $v = \sqrt{2kT/m}$, называемом наиболее вероятной скоростью. Название связано с тем, что вероятность наугад взятой молекулы газа иметь скорость в интервале заданного размера будет наибольшей, когда этот интервал расположен под максимумом кривой функции распределения, т. е. когда он содержит в себе наивероятнейшую скорость.

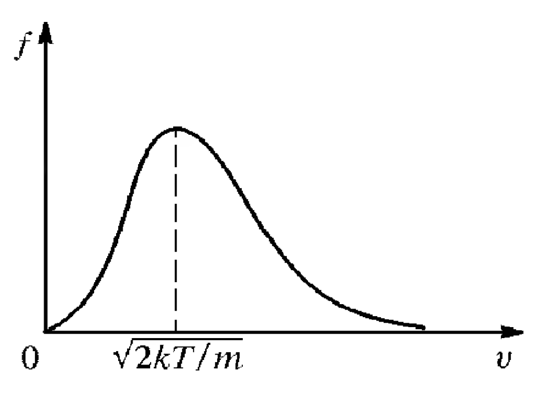


Рис. 74. График максвелловского распределения молекул газа по скоростям

Площадь, ограниченная графиком функции и осью v, дает полное число молекул в единице объема n.

Зависимость распределения по скоростям от температуры. При повышении температуры кривая распределения молекул по скоростям деформируется так, как показано на рис. 75. Максимум кривой смещается при увеличении T в область бóльших значений v. Максимальное значение $f(v)$ при этом убывает, так что площадь под кривой остается неизменной.

Приведенные выше функции распределения молекул газа по скоростям были впервые получены Максвеллом и носят его имя.

Рис. 75. Функции распределения молекул газа по скоростям при разных температурах

Экспериментальная кривая распределения Максвелла. Экспериментальное измерение скоростей молекул и проверка закона распределения Максвелла осуществляются различными методами, использующими молекулярные пучки. Один из первых таких опытов был проделан О. Штерном в 1920 г.

В экспериментальной установке Штерна использовались два коаксиальных цилиндра, скрепленных между собой (рис. 76). Цилин-

дры можно привести во вращение с определенной угловой скоростью. Вдоль оси O цилиндров натянута тонкая платиновая проволока, покрытая слоем серебра. При пропускании электрического тока проволока нагревается, и серебро начинает испаряться. Вся установка помещена в непрерывно откачиваемый вакуумный сосуд.

Испарившиеся атомы серебра летят прямолинейно. Пролетев сквозь прорезь A во внутреннем цилиндре, они оседают на охлаждаемой поверхности внешнего цилиндра, образуя на ней четкую полоску B металлического серебра (рис. 76*а*). Если цилиндры привести в быстрое вращение с постоянной угловой скоростью, то полоска осажденных атомов на поверхности внешнего цилиндра оказывается смещенной от прежнего положения на некоторое расстояние l (рис. 76*б*) и заметно размытой. Смещение вызвано тем, что за время пролета атомом расстояния от щели до внешнего цилиндра вся система успевает повернуться на некоторый угол φ. По смещению полоски можно судить о величине скорости атомов серебра. Размытие полоски происходит потому, что атомы серебра имеют различные скорости, и такой опыт в принципе дает возможность измерить распределение по скоростям атомов пучка.

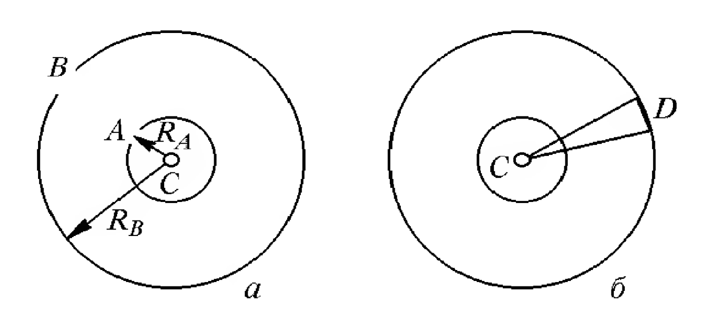


Рис. 76. Схема опытов Штерна

Впоследствии аналогичные опыты неоднократно воспроизводились в различных вариантах для разных веществ. В частности, такие опыты давали возможность выделять из пучка группы атомов, скорости которых лежали в определенном заданном интервале скоростей. Во всех случаях было получено хорошее согласие с законом Максвелла распределения молекул по скоростям.

Вычисление средних значений. Знание статистических функций распределения дает возможность вычислять средние значения микроскопических параметров, не зная их значений у отдельных молекул.

Например, с помощью функции распределения Больцмана можно найти высоту центра масс молекул газа в поле тяжести. Координата z_c центра масс N материальных точек одинаковой массы определяется формулой

$$z_c = \frac{1}{N}\sum_{i=1}^{N} z_i. \tag{16}$$

Нам неизвестны координаты z_i отдельных молекул, но функция распределения Больцмана говорит о том, сколько молекул имеют значения в интервале от z до $z + dz$:

$$dN = n(z)dz. \tag{17}$$

Теперь, при вычислении по формуле (16) можно сначала сгруппировать слагаемые z_i, лежащие в интервале от z до $z+dz$. Их вклад в сумму равен $zn(z)dz$. Затем остается просуммировать по всем таким группам, т. е. проинтегрировать по z от 0 до ∞:

$$z_c = \frac{1}{N} \int\limits_0^\infty zn(z)\,dz. \tag{18}$$

Подставляя сюда $n(z)$ из (8) и учитывая, что

$$N = \int\limits_0^\infty n(z)\,dz, \tag{19}$$

после вычислений получаем

$$z_c = \frac{kT}{mg}. \tag{20}$$

С помощью функции распределения Максвелла можно вычислить средние значения величин, зависящих от скоростей молекул. Для этого нужно разбить молекулы на группы с примерно одинаковыми значениями скоростей, лежащими в интервале от v до dv, найти вклад каждой такой группы в вычисляемое среднее, а затем просуммировать по всем группам молекул, т. е. проинтегрировать по v от 0 до ∞. Например, для нахождения среднего значения квадрата скорости нужно вычислить интеграл

$$\langle v^2 \rangle = \int\limits_0^\infty v^2 q(v)\,dv, \tag{21}$$

где $q(v)$ определяется формулой (14). Расчет дает значение

$$\langle v^2 \rangle = \frac{3kT}{m}, \tag{22}$$

что соответствует физическому смыслу температуры как меры средней кинетической энергии хаотического теплового поступательного движения молекул (см. формулы (7) и (12) предыдущего параграфа).

Для среднего значения абсолютной величины скорости, аналогично (21), можно получить

$$\langle v \rangle = \int\limits_0^\infty vq(v)\,dv = \sqrt{\frac{8kT}{\pi m}} = \sqrt{\frac{8RT}{\pi M}}. \tag{23}$$

Например, для молекул азота ($M = 28$ г/моль) при комнатной температуре получаем $\langle v \rangle = 470$ м/с.

- Как вы понимаете утверждение о хаотичности теплового движения? Что значит, что хаос имеет свои законы?

- Из какого физического условия выводится равновесное распределение молекул газа в поле тяжести? Где при выводе барометрической формулы используется предположение о постоянстве температуры?
- Почему барометрическая формула неприменима к зависимости от высоты давления в жидкости?
- Какой физический смысл имеет функция распределения молекул по координатам? по скоростям? Что общего имеют эти функции распределения?
- Объясните, почему график функции распределения по проекции скорости имеет симметричную форму, а кривая распределения молекул по модулю скорости асимметрична.
- Как объяснить, что среднее и наиболее вероятное значения проекции скорости на любое направление равны нулю, а соответствующие величины для модуля скорости отличны от нуля и не равны друг другу?
- Как, зная функцию распределения по проекции скорости $f(v_x)$, получить функцию распределения по модулю скорости?
- Что можно сказать о свойствах кривой, описывающей распределение молекул по скоростям? Какой смысл имеет ограничиваемая этой кривой площадь?
- Как изменяются с температурой положение максимума кривой функции распределения молекул по скоростям и его высота?
- Найдите количественную связь смещения полоски осажденных атомов в опыте Штерна с характерной скоростью молекул.

§ 23. Тепловое равновесие в статистической механике

Тепловое равновесие — это всегда динамическое равновесие. Тепловое движение атомов или молекул, из которых состоит макроскопическая система, никогда не прекращается. Поэтому макроскопические величины, характеризующие систему в целом, строго говоря, никогда не остаются постоянными, а испытывают малые беспорядочные колебания вблизи некоторых средних значений. Такие хаотические отклонения от средних значений тех или иных величин, происходящие в течение малых промежутков времени, называются *флуктуациями*. Относительная величина флуктуаций тем больше, чем меньше размеры изучаемой системы.

Флуктуации макроскопических параметров. Яркий пример флуктуаций — это дрожание зеркальца чувствительного гальванометра. Макроскопическая система, состоящая из подвижной катушки гальванометра, подвешенной на упругой кварцевой нити, в состоянии механического равновесия была бы совершенно неподвижной, если бы не тепловое движение. Удары молекул воздуха, совершающих тепловое движение, приводят к тому, что угол поворота зеркальца

испытывает хаотические колебания вблизи положения механического равновесия. Фактически это то же броуновское движение, которое отличается от рассмотренного выше движения взвешенной в жидкости частицы только тем, что здесь рассматривается не поступательное, а вращательное движение вблизи устойчивого, а не безразличного положения равновесия. Интенсивность такого движения зависит от температуры, оно принципиально неустранимо и ставит предел чувствительности измерительной аппаратуры.

Пространственное распределение молекул. Основные закономерности флуктуаций можно подметить, рассматривая пространственное распределение молекул идеального газа внутри сосуда в состоянии теплового равновесия. В среднем газ равномерно заполняет весь сосуд, т. е. концентрация молекул всюду одинакова. Разделим мысленно сосуд на две равные части. Пусть число молекул слева равно n_1, справа — n_2. Сумма $n_1 + n_2$ есть полное число молекул в сосуде.

В равновесии в среднем $n_1 = n_2$, но так как это равновесие динамическое, то в каждый момент времени вследствие хаотического движения это верно лишь приближенно, потому что молекулы непрерывно переходят из одной половины сосуда в другую и обратно. В принципе ничто не мешает им вообще в какой-то момент времени собраться в одной половине сосуда. Однако такое событие будет крайне маловероятным.

Чтобы убедиться в этом, рассмотрим, какими способами молекулы могут распределяться между половинами сосуда. Одну молекулу можно разместить в сосуде двумя способами — либо в левой, либо в правой его половине. Исходя из симметрии, естественно предположить, что вероятность найти молекулу в какой-то одной половине сосуда равна $1/2$, если, конечно, сосуд разделен на строго равные части. Две молекулы можно распределить в сосуде четырьмя (2^2) способами, ибо для каждого из двух способов распределения одной молекулы существует два способа распределения другой. Для трех молекул число способов распределения равно восьми (2^3), так как для каждого из четырех возможных вариантов распределения первых двух молекул есть две возможности распределения третьей молекулы, и т. д. В случае N молекул число способов распределения равно 2^N.

Вероятности распределений молекул по половинам сосуда. Будем считать, что вероятность нахождения любой наугад взятой молекулы в определенной половине сосуда не зависит от того, где в этот момент находятся все остальные молекулы. Это верно для идеального газа, молекулы которого не взаимодействуют между собой и имеют пренебрежимо малые размеры, так что их собственный объем значительно меньше объема сосуда. В этом случае каждый из

2^N способов распределения N молекул в сосуде имеет одну и ту же вероятность, равную $(1/2)^N$. Разъясним это подробнее.

Молекулы газа находятся в тепловом движении, и их расположение в сосуде непрерывно изменяется. Предположим, что мы можем делать мгновенные «фотографии» положения молекул в сосуде. На каждой такой «фотографии» мы увидим один из мыслимых способов распределения молекул. Распределения на любых двух снимках считаются одинаковыми, если в какой-либо половине сосуда, например левой, на обоих снимках находятся одни и те же молекулы. Утверждение о равной вероятности всех 2^N распределений означает, что какое-либо распределение будет встречаться на снимках не чаще и не реже других, в среднем один раз в каждой серии, содержащей 2^N «фотографий».

Разместить все N молекул в одной половине сосуда можно только одним способом. Поэтому вероятность того, что весь газ самопроизвольно соберется в одной половине сосуда, равна $(1/2)^N$. При больших N это очень малая величина. Например, для газа, содержащего всего 100 молекул, вероятность такого события равна $2^{-100} \approx 10^{-30}$. В среднем только на одном из 10^{30} снимков мы бы увидели, что одна из половин сосуда пуста. (Чтобы представить себе, насколько велико это число, вспомним, что, по преданию, изобретатель шахматной игры потребовал в награду «всего» $2^{64} \approx 10^{19}$ зерен пшеницы. Это оказалось больше, чем количество зерен, которое может уместиться в амбаре высотой 5 м и шириной 20 м, протянувшемся от Земли до Солнца.) И это для газа, состоящего всего из ста молекул! А для одного моля газа, где число молекул составляет $6 \cdot 10^{23}$, вероятность всем молекулам оказаться в одной половине сосуда настолько мала, что просто невозможно найти для нее подходящее сравнение в мире образов, доступных человеческому воображению.

Разместить N молекул в сосуде так, чтобы в определенной половине сосуда была только одна молекула, а в другой остальные $(N - 1)$, можно, очевидно, N способами, поскольку в первой половине сосуда может при этом находится любая из N молекул. Вероятность такого распределения молекул равна $N/2^N$.

Подсчитаем теперь вероятность того, что в определенной половине сосуда находится заданное число n молекул. Эта вероятность равна отношению числа способов распределения, при которых в выбранной половине сосуда находятся любые n молекул, к 2^N — полному числу возможных распределений.

Число таких распределений, очевидно, равно числу способов, которыми можно выбрать n молекул из совокупности N молекул, т. е. числу сочетаний из N по n. Это число сочетаний равно

$$C_N^n = \frac{N!}{n!(N-n)!}\,. \qquad (1)$$

Поэтому вероятность P_n того, что в какой-нибудь, например левой, половине сосуда находится n молекул, равна

$$P_n = \frac{N!}{2^N n!(N-n)!}. \qquad (2)$$

Такое распределение вероятностей называется биномиальным.

Вероятность равномерного распределения. Теперь с помощью (2) легко рассчитать, как часто на наших «фотографиях» будут встречаться те или иные распределения молекул. Для газа из ста молекул вероятность того, что молекулы распределятся между половинами сосуда строго поровну, приблизительно равна 1/18, т. е. такое распределение встречается в среднем один раз на каждые 18 снимков.

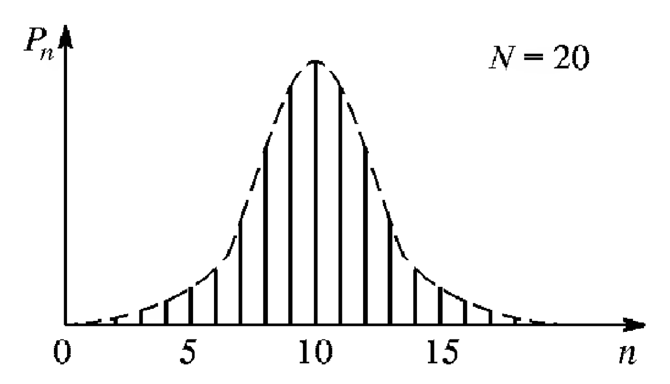


Рис. 77. Вероятность того, что в сосуде, содержащем 20 молекул газа, n молекул находятся в одной половине сосуда, а остальные $(20-n)$ — в другой

Итак, строго равномерное распределение молекул по объему встречается не так уж и часто. Но если мы подсчитаем вероятности таких распределений молекул, при которых в одной половине сосуда будет на две, четыре и вообще на небольшое число молекул больше, чем в другой, и просуммируем эти вероятности, то мы увидим, что вероятность появления любого из таких распределений будет близка к единице, а очень неравномерные распределения, при которых число молекул справа и слева различается сильно, будут встречаться крайне редко.

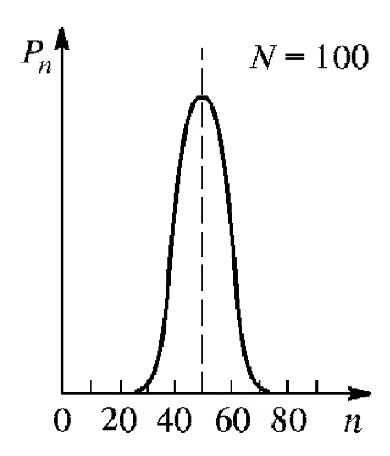


Рис. 78. Та же вероятность в случае, когда сосуд содержит 100 молекул газа

Представление о характере распределения вероятностей (2) дает рис. 77, построенный для газа из $N = 20$ молекул. По оси абсцисс отложено число молекул в одной (например, левой) половине сосуда, а по оси ординат — вероятности соответствующих распределений молекул. Приведенную зависимость вероятностей от n можно охарактеризовать плавной кривой — огибающей, которая показана штриховой линией на рис. 77. С увеличением полного числа молекул газа эта кривая становится все более и более острой. Для $N = 100$ она приведена на рис. 78.

Таким образом, знание функции распределения вероятностей позволяет не только вычислять средние значения макроскопических

параметров, но и находить характерные отклонения этих величин от своих средних значений, дающие представление о величине флуктуаций.

Общие закономерности флуктуаций. Теперь мы можем сделать следующий общий вывод: если полное число молекул в изучаемой системе велико, то флуктуации плотности, соответствующие существенно неоднородному распределению, почти никогда не возникают.

Описанные выше закономерности справедливы не только для флуктуаций плотности, но и для флуктуаций других макроскопических параметров. Если число частиц в системе очень велико, то относительная величина флуктуаций любого параметра очень мала по сравнению со средним значением этого параметра, и ею почти всегда можно пренебречь. Поэтому мы обычно не осознаем факта существования флуктуаций, когда имеем дело с большими макроскопическими системами. Но если система достаточно мала, то флуктуации могут быть легко обнаружены и часто имеют большое значение.

Как уже упоминалось, тепловые флуктуации ставят предел чувствительности любых измерительных приборов. При измерении малых значений физических величин или слабых сигналов флуктуации в чувствительном элементе измерительного прибора могут оказаться сравнимыми с измеряемой величиной. Флуктуации, возникающие в самом измерительном приборе, затрудняют измерения и поэтому называются шумами.

- Почему при хаотическом тепловом движении все молекулы макроскопического количества газа не могут самопроизвольно собраться в одной половине сосуда?
- Что такое флуктуации макроскопических параметров? Приведите примеры появления флуктуаций в физическом эксперименте.
- Какие соображения вы можете привести в пользу предположения о равной вероятности каждой молекуле идеального газа оказаться в одной из половин сосуда?
- Как подсчитать вероятность того, что в определенной половине сосуда находится n молекул газа из полного числа N?
- Чему равна сумма вероятностей P_n, даваемых формулой (2), по всем значениям n от 0 до N?
- Как изменяется вид функции распределения этих вероятностей с увеличением полного числа N молекул газа?
- Каковы максимальные значения вероятностей P_n на рисунках 77 и 78?
- Как относительная флуктуация макроскопических параметров зависит от размеров системы, т. е. от числа молекул N?

△ **Статистический подход.** На приведенном выше примере расчета вероятностей распределения молекул идеального газа по половинам сосуда можно проиллюстрировать основные идеи Гиббса построения равновесной статистической механики. При этом все макроскопические характеристики рассматриваемой молекулярной модели термодинамической системы должны вводиться без ссылок на термодинамику, и только затем можно найти их термодинамические эквиваленты.

Основное понятие статистического метода — это микроскопическое состояние. Задать его — значит задать механическое состояние системы, т. е. координаты и скорости всех молекул в один момент времени. Но в некоторых случаях достаточно в определение микросостояния включать только часть характеристик отдельных молекул. Так, в приведенном примере мы под микросостоянием фактически понимали только информацию о том, в какой половине сосуда находится каждая молекула. Под макросостоянием в этом случае понимается информация о том, сколько молекул находится в левой, а тем самым и сколько в правой половине сосуда.

Равные вероятности микросостояний. Основной гипотезой в методе Гиббса является предположение о равной вероятности реализации любых возможных микроскопических состояний замкнутой системы, для которой зафиксированы энергия, объем и число частиц. Именно это предположение в рассмотренном выше примере приводило к биномиальному распределению вероятностей P_n. Гипотеза равной вероятности различных микросостояний считается справедливой не только для идеального газа, но и для любых сколь угодно сложных термодинамических систем.

Статистический вес макросостояния. Вероятности реализации различных макросостояний при этом, в отличие от микросостояний, уже не равны друг другу, поскольку каждое макросостояние может быть реализовано своим вполне определенным числом микросостояний. Так, макросостояние, в котором все молекулы находятся в одной (левой) половине сосуда, реализуется единственным микросостоянием, а макросостояние, когда в левой половине находится одна из N молекул, реализуется N различными микросостояниями.

Статистическим весом или *термодинамической вероятностью* некоторого макросостояния системы называется число различных микросостояний, которыми может быть реализовано данное макросостояние. В рассмотренном примере статистический вес макросостояния с определенным числом n молекул в левой половине — это число способов такого распределения N молекул газа по половинам сосуда. Оно равно числу сочетаний C_N^n.

Как определить понятие теплового равновесия при статистическом подходе? Рассмотрим две термодинамические системы, находящиеся в тепловом контакте и образующие вместе замкнутую систему с определенной энергией E. Число частиц и объем обеих подсистем фиксированы, но они могут обмениваться энергией в форме теплопередачи. Пусть на долю первой подсистемы приходится энергия E_1, а на долю второй — остальная энергия $E - E_1$. Число микросостояний всей замкнутой системы $W(E_1|E)$ равно при этом произведению числа различных микросостояний первой подсистемы $W_1(E_1)$ на число микросостояний второй подсистемы $W_2(E - E_1)$:

$$W(E_1|E) = W_1(E_1) \cdot W_2(E - E_1). \qquad (3)$$

Значение $W(E_1|E)$ будет разным в зависимости от того, какая часть E_1 полной энергии E приходится на первую подсистему.

Тепловое равновесие между подсистемами определяется как макроскопическое состояние с *наибольшим статистическим весом*, т. е. состояние с таким распределением энергии между подсистемами, при котором $W(E_1|E)$ как функция E_1 имеет максимальное значение. В разобранном примере состоянию равновесия соответствовало распределение молекул между половинами сосуда поровну.

Найдем, при каком условии статистический вес $W(E_1|E)$ максимален. Приравняем нулю производную $dW(E_1|E)/dE_1$:

$$\frac{dW_1(E_1)}{dE_1} W_2 + W_1 \frac{dW_2(E - E_1)}{dE_1} = 0. \qquad (4)$$

Производная от $W_2(E - E_1)$ по E_1 равна взятой с противоположным знаком производной от W_2 по ее аргументу $E_2 = E - E_1$. Поэтому (4) можно записать в виде

$$\frac{1}{W_1}\frac{dW_1}{dE_1} = \frac{1}{W_2}\frac{dW_2}{dE_2}. \qquad (5)$$

Равенство (5), выражающее условие максимума статистического веса, равносильно равенству производных от логарифмов:

$$\frac{d \ln W_1}{dE_1} = \frac{d \ln W_2}{dE_2}. \qquad (6)$$

Энтропия и температура. В статистической механике энтропия S и термодинамическая температура T определяются следующими равенствами:

$$S = k \ln W, \qquad (7)$$

$$\frac{1}{T} = \frac{dS}{dE} = k \frac{d \ln W}{dE}. \qquad (8)$$

При вычислении производной в (8) объем системы V и число частиц в ней N следует считать фиксированными параметрами.

Определение температуры (8) обусловлено тем, что в состоянии теплового равновесия, характеризующимся равенством температур, именно такая величина, как видно из (6), одинакова у обеих подсистем.

Определенная равенством (7) энтропия как характеристика макросостояния термодинамической системы обладает следующими свойствами.

Во-первых, энтропия аддитивна: энтропия всей системы равна сумме энтропий составляющих ее частей. Это непосредственно следует из равенства (3).

Во-вторых, энтропия системы максимальна в состоянии теплового равновесия. Это значит, что введенная определением (7) энтропия обладает таким же свойством, что и энтропия, введенная в термодинамике: согласно второму закону термодинамики энтропия замкнутой системы имеет максимальное значение в состоянии теплового равновесия.

Покажем согласованность определений энтропии (7) и температуры (8) с определениями этих величин в термодинамике. Пусть два тела с разными температурами T_1 и T_2 приводятся в тепловой контакт. В результате теплообмена они приходят в состояние теплового равновесия. Энтропия S всей системы при этом увеличивается. Запишем выражение для изменения S:

$$\Delta S = \Delta S_1 + \Delta S_2 = \frac{dS_1}{dE_1}\,\Delta E_1 + \frac{dS_2}{dE_2}\,\Delta E_2 > 0. \qquad (9)$$

Пусть некоторое количество теплоты ΔE переходит от тела с температурой T_2 к телу с температурой T_1. При этом энергия E_1 первого тела возрастает, а энергия E_2 убывает: $\Delta E_1 = -\Delta E_2 = \Delta E > 0$. В соответствии с определением (8) производные в (9) равны обратным температурам тел:

$$\frac{dS_1}{dE_1} = \frac{1}{T_1}, \qquad \frac{dS_2}{dE_2} = \frac{1}{T_2}.$$

Теперь (9) переписывается в виде

$$\Delta S = \Delta E\left(\frac{1}{T_1} - \frac{1}{T_2}\right) > 0.$$

Отсюда следует, что $T_2 > T_1$, т. е. что при теплообмене энергия переходит от тела с более высокой температурой к телу с более низкой температурой.

Задача о расширении газа в пустоту. Покажем, что статистическое определение (7) энтропии S приводит при решении задачи о

необратимом расширении газа при разрушении перегородки в сосуде (см. § 19) к тому же результату, что и термодинамическое определение. Напомним условие задачи.

Идеальный газ находится в одной половине теплоизолированного сосуда, разделенного перегородкой на две равные части. Определить изменение энтропии газа, если перегородка внезапно разрушается и газ заполнит весь сосуд.

Статистический вес начального состояния газа равен единице: $W_1 = 1$, так как существует лишь одно микросостояние, при котором все молекулы находятся в одной половине сосуда. В конечном состоянии молекулы распределены по половинам сосуда почти поровну (строго поровну, если отвлечься от флуктуаций). Вследствие исключительной остроты максимума кривой распределения вероятностей различных распределений молекул газа по половинам сосуда приближенно можно считать, что все остальные $(2^N - 1)$ микросостояний системы соответствуют конечному макросостоянию:

$$W_2 \approx 2^N - 1 \approx 2^N.$$

Поэтому, в соответствии с (7)

$$\Delta S = k(\ln W_2 - \ln W_1) = kN \ln 2.$$

При строгом решении следует воспользоваться формулой (1) при $n = N/2$:

$$W_2 = C_N^{N/2} = \frac{N!}{[(n/2)!]^2}, \tag{10}$$

и формулой Стирлинга

$$n! \approx (2\pi n)^{1/2} n^n.$$

В результате имеем

$$\ln W_2 \approx N \ln 2 - \frac{1}{2} \ln N.$$

Поскольку $N \approx 10^{23}$, то $\ln N \approx 50$, и для энтропии получаем прежнее значение

$$\Delta S = kN \ln 2. \tag{11}$$

Распределение Гиббса. Гипотеза Гиббса о равной вероятности всех микросостояний замкнутой системы позволяет получить функцию распределения вероятностей для любой малой подсистемы, находящейся в тепловом контакте с очень большой системой — термостатом.

Рассмотрим замкнутую систему, состоящую из рассматриваемой подсистемы и термостата. Пусть E_0 — энергия всей замкнутой системы, а E — энергия выделенной подсистемы. Тогда на долю термостата приходится энергия $E_0 - E$. Очевидно, что вероятность

$P(E)$ подсистеме иметь энергию E пропорциональна числу микросостояний $W(E_0 - E)$ термостата с энергией $E_0 - E$. Поэтому для отношения двух вероятностей $P(E_1)$ и $P(E_2)$ можно написать

$$\frac{P(E_1)}{P(E_2)} = \frac{W(E_0 - E_1)}{W(E_0 - E_2)}.$$

Используя определение энтропии (7), перепишем это отношение в виде

$$\frac{P(E_1)}{P(E_2)} = \exp\frac{\Delta S}{k}, \tag{12}$$

где $\Delta S = S(E_0 - E_1) - S(E_0 - E_2)$. Учитывая, что энергии подсистемы E_1 и E_2 малы по сравнению с энергией термостата, выражение для ΔS можно приближенно записать следующим образом:

$$\Delta S = -\frac{dS}{dE}E_1 + \frac{dS}{dE}E_2,$$

или, с учетом (8),

$$\Delta S = \frac{1}{T}(E_2 - E_1).$$

Подставляя ΔS в (12), находим

$$\frac{P(E_1)}{P(E_2)} = \exp\frac{E_2 - E_1}{kT} = \frac{\exp(-E_1/kT)}{\exp(-E_2/kT)}.$$

Итак, мы получаем, что вероятность того, что малая подсистема, находящаяся в тепловом контакте с термостатом при температуре T, будет иметь энергию E, пропорциональна экспоненте $\exp(-E/kT)$:

$$P(E) \sim \exp\left(-\frac{E}{kT}\right). \tag{13}$$

Формула (13) дает распределение вероятностей для любой физической величины, характеризующей подсистему. При этом в показателе экспоненты играет роль лишь та часть энергии подсистемы, которая зависит от этой величины.

Если, например, в качестве подсистемы взять одну молекулу, то (13) дает распределение для ее координат, скорости, магнитного или электрического момента, и т. д. В частности, рассматривая распределение вероятностей для проекции скорости v_x, нужно взять в качестве E ту часть кинетической энергии, которая зависит от v_x, т. е. $\frac{1}{2}mv_x^2$. В результате сразу приходим к формуле (12) § 22. При этом коэффициент пропорциональности перед экспонентой получается из условия нормировки: сумма всех вероятностей должна быть равна единице. ▲

- Приведите соображения, обосновывающие выражение (3) для числа микросостояний составной замкнутой системы. При каких условиях это выражение справедливо?
- Как понимается условие теплового равновесия в статистической механике?
- Объясните, почему состоянию теплового равновесия идеального газа в сосуде соответствует распределение молекул поровну между половинами сосуда.
- Как определяются энергия и температура в статистической механике на основе представления о статистическом весе макроскопического состояния системы?
- Поясните, почему энтропия S, определяемая равенством (7), увеличивается при переходе замкнутой системы в состояние термодинамического равновесия.
- Приведите доводы, свидетельствующие о том, что статистические определения энтропии (7) и температуры (8) согласуются с определениями этих величин в термодинамике.
- Поясните, почему в приведенном выводе распределения Гиббса считалось, что вероятность подсистеме иметь энергию E пропорциональна числу состояний термостата с энергией $E_0 - E$.
- Где в выводе распределения Гиббса использовалось предположение о том, что выделенная подсистема много меньше термостата?
- Покажите, как из общего распределения Гиббса (13) получить барометрическую формулу?

§ 24. Статистическая природа необратимости тепловых процессов

Термодинамический подход не позволяет вскрыть внутреннюю природу необратимости реальных процессов в макроскопических системах. Опираясь на эксперимент, он только фиксирует факт необратимости (второй закон термодинамики). Молекулярно-кинетический подход позволяет проанализировать причины такой необратимости реальных процессов и определенной направленности энергетических превращений в природе.

Гипотетический вечный двигатель. Рассмотрим с точки зрения молекулярно-кинетической теории модель гипотетического «вечного» двигателя второго рода, изображенную на рис. 79. Предположим, что этот вечный двигатель работает следующим образом: газ самопроизвольно собирается в левой половине цилиндра, после чего поршень подвигают вплотную к газу. При таком перемещении внешние силы работы не совершают, так как собравшийся в левой половине газ не ока-

зывает давления на поршень. Затем подводим к газу теплоту и заставляем его изотермически расширяться до прежнего объема. При этом газ совершает работу за счет подводимой теплоты. После того, как поршень перейдет в крайнее правое положение, будем ждать, пока газ снова не соберется самопроизвольно в левой половине сосуда, и затем повторяем все снова. В результате получилась периодически действующая машина, которая совершает работу только за счет получения теплоты от окружающей среды.

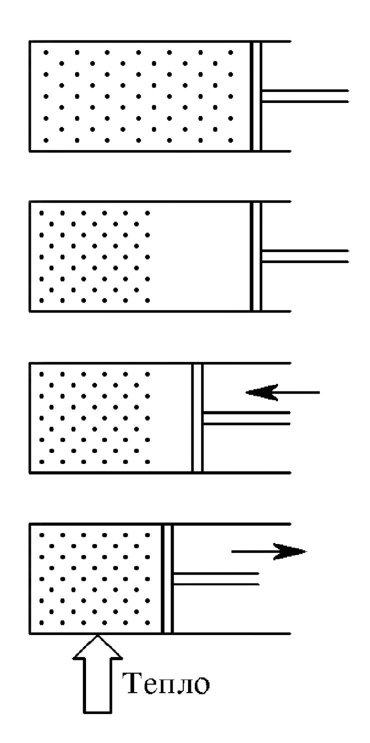


Рис. 79. Один из вариантов «вечного» двигателя второго рода

Молекулярно-кинетическая теория позволяет сразу объяснить, почему такое устройство не будет работать. Как мы видели, вероятность того, что газ, содержащий большое число молекул, хотя бы один раз самопроизвольно соберется в одной половине сосуда, ничтожно мала. И уж совершенно невозможно себе представить, чтобы это могло повторяться раз за разом по мере работы машины.

О необратимых процессах. Теперь можно указать, какой смысл вкладывается в понятие необратимого процесса: процесс является необратимым, если обратный процесс в действительности почти никогда не происходит. Строго запрета для такого процесса нет — он просто слишком маловероятен, чтобы его можно было наблюдать на опыте. Так, рассмотренный пример вечного двигателя второго рода основывался на предположении о возможности самопроизвольного сосредоточения газа в одной половине сосуда. Такой процесс является обратным для процесса расширения газа в пустоту. Расширение газа в пустоту представляет собой один из наиболее ярких примеров необратимых процессов — обратный процесс в макроскопической системе никогда не наблюдался.

Таким образом, с точки зрения представлений статистической механики второй закон термодинамики утверждает, что в природе в макроскопических системах процессы развиваются в таком направлении, когда менее вероятные состояния системы заменяются на более вероятные. Такая интерпретация второго закона термодинамики была впервые предложена Больцманом.

При рассмотрении флуктуаций плотности идеального газа было выяснено, что состояния газа, при которых распределение молекул близко к равномерному, встречаются гораздо чаще, чем далекие от равновесия состояния с сильно неравномерным распределением молекул. Другими словами, состояния с неравномерным распределением молекул по объему, при которых число молекул в правой и левой половинах сосуда сильно различаются, имеют гораздо меньшую вероятность, чем состоя-

ния с почти равномерным распределением, близким к равновесному. Итак, необратимый процесс приближения к равновесию — это переход к наиболее вероятному макроскопическому состоянию.

Необратимые процессы и разрушение порядка. Сказанное выше о природе необратимости реальных процессов можно сформулировать и несколько иначе. Можно сказать, что необратимый переход к равновесию — это переход от в сильной степени упорядоченных неравновесных состояний к менее упорядоченным, хаотическим состояниям.

При расширении газа в пустоту начальное состояние, когда газ занимает часть предоставленного ему объема, является в значительной мере упорядоченным, в то время как конечное состояние теплового равновесия, когда газ равномерно распределен по всему объему сосуда, является совершенно неупорядоченным.

Другой пример — направленный пучок молекул газа, входящий в откачанный сосуд. Установление равновесного максвелловского распределения молекул по скоростям представляет собой необратимый процесс перехода системы из упорядоченного состояния, когда все молекулы имеют почти одинаковые по модулю и направлению скорости, в конечное состояние, характеризующееся полной хаотичностью движения молекул.

С этой точки зрения легко понять устанавливаемую вторым законом термодинамики определенную направленность энергетических превращений в замкнутой системе. Когда тело получает некоторое количество теплоты за счет совершения механической работы, то это означает необратимое превращение кинетической энергии упорядоченного макроскопического движения в кинетическую энергию хаотического движения молекул. Превращение теплоты в работу, наоборот, означает превращение энергии беспорядочного движения молекул в энергию упорядоченного движения макроскопического тела — такой самопроизвольный переход, как мы видели, в принципе возможен, но исключительно маловероятен.

Флуктуации как отклонения от второго закона термодинамики. Необратимый характер процессов перехода в состояние теплового равновесия, устанавливаемый вторым законом термодинамики, справедлив только для больших макроскопических систем. С термодинамической точки зрения изолированная система, пришедшая в состояние теплового равновесия, не может самопроизвольно выйти из этого состояния. Однако статистическая механика допускает существование флуктуаций, которые фактически представляют собой самопроизвольные отклонения системы от равновесия.

Как уже отмечалось, чем больше частиц в системе, тем меньше относительная величина флуктуаций любого макроскопического параметра, и для достаточно большой системы флуктуациями вообще можно пренебречь. Именно поэтому для таких систем справедлив

второй закон термодинамики, в котором утверждается возрастание энтропии в замкнутых системах. При статистическом определении энтропии второй закон утрачивает абсолютный характер и превращается в статистический закон: за каким-либо заданным состоянием замкнутой системы будут следовать состояния, более вероятные если не с необходимостью, то в подавляющем большинстве случаев.

В системах с небольшим числом частиц относительная величина флуктуаций велика, т. е. самопроизвольные отклонения какой-либо величины от ее среднего значения могут быть сравнимы с самим средним значением. Такая система часто самопроизвольно выходит из состояния равновесия, и второй закон термодинамики здесь неприменим. Характерный пример нарушения второго закона термодинамики в достаточно малых системах — броуновское движение, при котором взвешенная в жидкости макроскопическая частица получает кинетическую энергию от молекул окружающей среды, хотя температура среды не выше, чем температура самой броуновской частицы.

- Как статистическая механика объясняет необратимость реальных тепловых процессов?
- Приведите примеры явлений, в которых наблюдается самопроизвольный выход системы из состояния термодинамического равновесия.
- Почему упорядоченные состояния характеризуются меньшей вероятностью по сравнению с неупорядоченными?

△ **Статистическая гипотеза.** Неизбежность тепловых процессов в природе приводит к тому, что статистическая механика систем многих частиц не исчерпывается законами обычной механики (хотя и опирается на них), а требует обязательного введения дополнительной *статистической гипотезы* в той или иной форме, например в виде предположения о равной вероятности различных микросостояний замкнутой системы.

Но в тех случаях, когда тепловые процессы оказываются несущественными, определенную информацию о свойствах термодинамической системы можно получить, опираясь только на механические представления. Тепловые процессы практически отсутствуют в условиях тепловой изоляции при наличии механического равновесия. В этих случаях протекающие явления обратимы и можно использовать модель адиабатического процесса.

Покажем, например, как можно получить уравнение адиабаты для одноатомного идеального газа, основываясь на существовании адиабатических инвариантов в механических системах. Напомним (см. кн. 1), что адиабатическим инвариантом называется характеризующая механическую систему величина, сохраняющаяся при медленном изменении внешних параметров. В частности, для шарика, упруго отражающегося от двух параллельных стенок, которые медленно сближаются или раздвигаются, адиабатическим ин-

вариантом является произведение расстояния между стенками на модуль скорости шарика.

В механической модели идеального газа как совокупности одноатомных молекул, упруго отражающихся от стенок сферического сосуда, адиабатическим инвариантом при медленном изменении объема V сосуда будет произведение характерного линейного размера (радиуса) R сосуда на модуль v скорости молекулы. В отсутствие теплообмена такая механическая модель адекватно описывает реальный адиабатический процесс сжатия или расширения. При этом сохраняет свой смысл и указанный адиабатический инвариант: $Rv = \text{const}$. Поскольку радиус R пропорционален $\sqrt[3]{V}$, а модуль скорости v пропорционален $\sqrt{T}$, то

$$\sqrt[3]{V}\,\sqrt{T} = \text{const},$$

или

$$TV^{2/3} = \text{const}.$$

Легко видеть, что это совпадает с полученным выше уравнением адиабаты в переменных T и V для одноатомного идеального газа, поскольку в этом случае $\gamma = C_p/C_V = 5/3$. ▲

- При каких условиях к системам из большого числа частиц применимы чисто механические представления, не опирающиеся на статистическую гипотезу?
- Получите уравнение адиабаты идеального газа, рассматривая сосуд цилиндрической формы, объем которого изменяется при медленном перемещении поршня.

§ 25. Газы, жидкости, фазовые переходы

Статистическая механика позволяет объяснить на основе некоторых моделей не только свойства вещества в газообразном состоянии, но и процессы перехода вещества из одного состояния в другое, в том числе фазовые превращения газ — жидкость, жидкость — твердое тело и т. д.

Фазы и компоненты термодинамической системы. *Фазой* вещества называется его макроскопическая однородная часть, отделенная от других частей границами раздела. Например, в закрытом сосуде над некоторым количеством воды находится смесь воздуха с водяными парами. Это *двухфазная* система, состоящая из двух фаз: жидкой (вода) и газообразной (смесь воздуха и паров воды). Если бы воздуха в сосуде не было, то в системе все равно было бы две фазы — жидкая (вода) и газообразная (пары воды), так что разделение вещества на фазы возможно и в однокомпонентной системе, а не только в смеси разных веществ.

Если к воде добавить спирт, то число фаз в системе не изменится, так как при разведении спирта в воде образуется физически однородная жидкость. Если же к воде добавить растительное масло, то получится система с двумя жидкими фазами — маслом и водой.

При подсчете числа фаз не имеет значения, сосредоточена ли определенная фаза в одном месте или состоит из нескольких частей, отделенных одна от другой другими фазами. Так, капельки тумана в воздухе образуют с ним двухфазную систему. Смесь газов, будучи многокомпонентной системой, всегда образует однофазную систему.

Фазовые превращения. Примерами фазовых превращений могут служить изменения *агрегатного состояния* вещества. Под агрегатными состояниями понимают твердое, жидкое и газообразное состояния вещества. Иногда в качестве четвертого агрегатного состояния вещества выделяют плазму. Твердое и жидкое состояния вещества называют *конденсированными*.

Испарением или *парообразованием* называют переход вещества из конденсированного состояния в газообразное. Обратный переход называется *конденсацией*. Переход из твердого состояния в жидкое называется *плавлением*, а обратный переход *затвердеванием* или *кристаллизацией*.

Твердое состояние одного и того же вещества может реализовываться в нескольких кристаллических модификациях. Превращение модификаций друг в друга тоже дает пример фазового перехода.

Равновесие фаз. При каких условиях различные фазы могут находиться в равновесии? Для механического равновесия давление по разные стороны границы соприкасающихся фаз должно быть одинаковым. Для теплового равновесия фазы должны иметь одинаковую температуру.

Равновесие фаз — это всегда динамическое равновесие, поскольку молекулы, совершая хаотическое тепловое движение, непрерывно переходят через границу из одной фазы в другую и обратно. Но эти потоки компенсируют друг друга.

Испарение и конденсация. Рассмотрим подробнее один из простейших примеров фазовых превращений — испарение жидкости и конденсацию пара.

Количество жидкости, налитой в открытый сосуд, постепенно уменьшается. При этом покидают жидкость преимущественно быстрые молекулы, кинетическая энергия которых достаточна для преодоления сил притяжения со стороны других молекул жидкости в приповерхностном слое. В результате внутренняя энергия жидкости уменьшается, и если не подводить к ней теплоту, то в процессе испарения ее температура будет понижаться.

Одновременно с испарением идет и обратный процесс конденсации, т. е. возвращение части молекул из пара в жидкость. В откры-

том сосуде испарение обычно не компенсируется конденсацией и количество жидкости уменьшается.

Если сосуд плотно закрыть, то вскоре установится динамическое равновесие и уровень жидкости не будет изменяться. Пар, находящийся в динамическом равновесии со своей жидкостью, называют *насыщенным паром*. Название подчеркивает, что в данном объеме при данной температуре не может находиться большее количество пара.

Наличие других газов или паров над поверхностью жидкости не влияет на процесс образования пара. Максимальное давление пара данного вещества, т. е. его максимальное парциальное давление, равно давлению насыщенного пара соответствующей жидкости и не зависит от присутствия других газов.

Давление насыщенного пара. Будем постепенно уменьшать объем пара над жидкостью, вдвигая закрывающий сосуд поршень. При этом часть вещества перейдет из газообразной фазы в жидкость, но давление пара не изменится, если поддерживается прежняя температура. Это значит, что давление насыщенного пара зависит только от температуры. Такой опытный факт можно качественно объяснить на основе представлений статистической механики.

Число вылетающих из жидкости молекул через участок границы единичной площади не зависит от того, сколько пара находится над жидкостью. В динамическом равновесии число вылетающих и влетающих молекул одинаково. Значит, и число молекул, влетающих в жидкость через единичную поверхность, не зависит от занимаемого паром объема. Так как это число при данной температуре определяется концентрацией n молекул пара, то концентрация, а тем самым и давление $p = nkT$, не зависят от объема, а определяются только температурой. Для каждого вещества давление насыщенного пара при данной температуре имеет свое вполне определенное значение.

С увеличением температуры давление насыщенного пара растет быстрее, чем давление идеального газа. Так происходит потому, что с увеличением температуры давление насыщенного пара растет не только вследствие увеличения средней кинетической энергии молекул, но и вследствие увеличения концентрации n молекул, т. е. плотности пара, за счет перехода части вещества из жидкой фазы в газообразную. Именно увеличение концентрации молекул насыщенного пара и является главной причиной роста давления с температурой.

Когда вся жидкость в закрытом сосуде испарится, пар при дальнейшем нагревании перестанет быть насыщенным, и его давление при постоянном объеме будет возрастать пропорционально термодинамической температуре, как и у идеального газа.

Кипение. По мере увеличения температуры интенсивность испарения жидкости в открытом сосуде увеличивается. Наконец, при некоторой температуре жидкость начинает кипеть: по всему ее объему

образуются быстро растущие пузырьки пара, поднимающиеся на поверхность. Это значит, что при кипении испарение жидкости происходит не только с ее открытой поверхности, но и внутрь пузырьков. Ясно, что рост пузырьков за счет испарения в них возможен лишь при такой температуре, когда давление внутри пузырька, т. е. давление насыщенного пара, становится равным давлению жидкости на данной глубине. Поэтому температура кипения жидкости зависит от внешнего давления. На высоте 7 км над уровнем моря температура кипения воды составляет примерно 70 °С. Наоборот, температура кипения в герметически закрытых автоклавах, где поддерживается повышенное давление, значительно выше 100 °С.

Интересными особенностями обладает кипение на границе несмешивающихся жидкостей. Их можно продемонстрировать следующим простым опытом. В стеклянный сосуд наливают некоторое количество четыреххлористого углерода (CCl_4), а сверху слой воды. При нормальном атмосферном давлении вода кипит при 100 °С, а четыреххлористый углерод — при 76,7 °С. Если медленно нагревать сосуд в водяной бане, то но границе раздела этих несмешивающихся жидкостей кипение начинается при 65,5 °С! Как объяснить это явление?

Кипение наступает, когда давление насыщенного пара сравняется с давлением в жидкости на той глубине, где образуется пузырек. Давление жидкости складывается из атмосферного давления и гидростатического давления столба жидкости. Если высота столба жидкости в сосуде несколько сантиметров, то гидростатическое давление составляет несколько тысячных от нормального атмосферного, и его можно не принимать во внимание. Давление насыщенных паров жидкости определяется ее температурой. В воде пузырьки содержат только пары воды, при 100 °С давление насыщенного пара воды равно нормальному атмосферному. В четыреххлористом углероде пузырьки содержат только пары CCl_4 и давление насыщенных паров равно атмосферному при 76,7 °С.

На границе раздела этих жидкостей пузырьки содержат как пары CCl_4, так и пары воды. Давление в этих пузырьках, на основании закона Дальтона, равно сумме парциальных давлений паров воды и CCl_4. Поэтому давление, равное атмосферному, устанавливается в пузырьках, находящихся на границе, при температуре, меньшей 76,7 °С. При этой температуре и начинается кипение на границе раздела. Опыт показывает, что это происходит при температуре 65,5 °С. Эту температуру можно установить без опыта, подбирая такую температуру по таблицам зависимости давления насыщенных паров воды и CCl_4 от температуры, при которой сумма давлений равна атмосферному. При 65,5 °С давление паров воды составляет 192 мм рт. ст., а давление паров CCl_4 — 568 мм рт. ст. Используя эти цифры, можно сделать вывод, что

в каждом пузырьке, образовавшемся на границе раздела, молекул CCl_4 будет почти в три раза больше, чем молекул воды. А так как и масса молекулы четыреххлористого углерода в девять раз больше массы молекулы воды, то его испарение происходит почти в 25 раз быстрее, чем испарение воды. Наблюдая кипение в пограничном слое в течение некоторого времени, можно убедиться, что нижний слой CCl_4 выкипает значительно быстрее, чем верхний слой воды.

Центры парообразования. Для процесса кипения необходимо, чтобы в жидкости существовали неоднородности — зародыши газообразной фазы, играющие роль центров парообразования. Обычно в жидкости присутствуют растворенные газы, которые выделяются пузырьками на дне и стенках сосуда и на взвешенных в жидкости пылинках. При нагревании эти пузырьки увеличиваются как за счет уменьшения растворимости газов с температурой, так и за счет испарения в них жидкости. Увеличившиеся в объеме пузырьки всплывают под действием архимедовой выталкивающей силы. Если верхние слои жидкости имеют более низкую температуру, то вследствие конденсации пара давление в них резко падает, и пузырьки «захлопываются» с характерным шумом. По мере прогревания всей жидкости до температуры кипения пузырьки перестают захлопываться и всплывают на поверхность: вся жидкость закипает.

Перегретая жидкость. Жидкость можно нагреть и до температуры, превышающей температуру кипения при данном давлении. Для этого ее нужно тщательно очистить от растворенных газов и механических примесей. Центры парообразования в такой жидкости практически отсутствуют, и кипение не начинается даже тогда, когда температура заметно превышает температуру кипения. Такая жидкость называется перегретой. Это неустойчивое состояние. Перегретая жидкость бурно закипает при появлении в ней центров парообразования.

Явление быстрого закипания перегретой жидкости используется в *пузырьковой камере* — современном приборе для регистрации частиц высоких энергий. Прохождение заряженной частицы через перегретую жидкость приводит к образованию вдоль следа частицы зародышей центров кипения. Образующиеся на зародышах пузырьки достигают размеров в доли миллиметра и могут быть сфотографированы при боковом освещении импульсным источником света. Изучение треков в пузырьковых камерах позволило открыть и исследовать много новых элементарных частиц.

Критическая температура. При увеличении температуры одновременно с ростом плотности насыщенного пара происходит уменьшение плотности жидкости вследствие ее теплового расширения. Если на одном графике изобразить зависимости плотности жидкости и ее насы-

щенного пара от температуры, то при некоторой температуре $T_{кр}$, называемой *критической*, обе кривые сливаются (рис. 80): плотность жидкости становится равной плотности насыщенного пара. При критической температуре исчезает различие не только в плотности, но и во всех других физических свойствах жидкости и насыщенного пара.

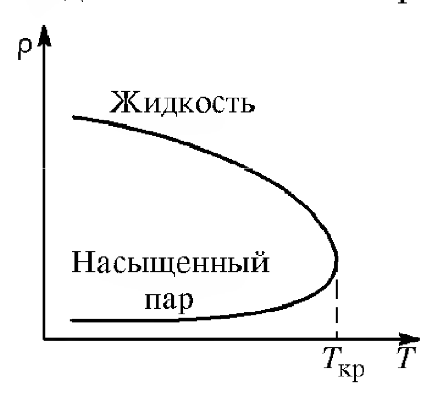


Рис. 80. Зависимость плотности жидкости и ее насыщенного пара от температуры

Понятие критической температуры было введено Д. И. Менделеевым. Каждое вещество характеризуется своей критической температурой. Например, у воды это 374,15 °С, а у окиси углерода CO_2 — около 31 °С.

Различие между газом и паром. Существование критической температуры позволяет провести различие между газом и паром. Газ можно перевести в жидкое состояние путем сжатия только при температуре ниже критической. Поэтому вещество в газообразном состоянии при температуре ниже критической и называют паром.

Таким образом, никакой принципиальной разницы между паром и газом нет. Это деление условно и имеет смысл потому, что мы привыкли к существованию в довольно узком интервале температур. Поэтому газом обычно называют вещество, давление насыщенного пара которого при обычных температурах выше атмосферного (например, углекислый газ). Напротив, используют термин «пар», когда при комнатной температуре давление насыщенного пара вещества меньше атмосферного давления (например, водяной пар).

Теплота испарения и конденсации. Фазовые переходы, такие как испарение и конденсация, всегда сопровождаются поглощением или выделением теплоты. Количество теплоты, которое необходимо сообщить веществу для того, чтобы превратить его из жидкого состояния в газообразное при постоянной температуре и постоянном давлении, называется *теплотой испарения* (или теплотой парообразования). Такое же количество теплоты выделяется при конденсации пара в жидкость при тех же условиях. Различают *удельную* теплоту испарения (измеряется в джоулях на килограмм) и *молярную* (измеряется в джоулях на моль).

Теплота испарения зависит от температуры. С ростом температуры она уменьшается и при критической температуре обращается в нуль.

Теплота парообразования представляет собой частный случай *теплоты фазового перехода*. Другой пример — *теплота плавления*, поглощаемая при плавлении и выделяющаяся при кристаллизации.

Влажность воздуха. Водяной пар в земной атмосфере, несмотря на огромные открытые водные поверхности, не является насыщенным.

Содержание водяного пара в воздухе характеризуется влажностью воздуха. От влажности воздуха во многом зависит состояние биосферы, самочувствие человека, работоспособность технических устройств, сохранность зданий и произведений искусства.

Влажность воздуха характеризуют парциальным давлением паров воды, которое иногда называют *абсолютной влажностью*. Однако значение абсолютной влажности еще не позволяет судить о том, насколько водяной пар в данных условиях далек от насыщения. Но именно от этого зависит интенсивность испарения воды с открытых поверхностей. Чтобы характеризовать близость состояния водяных паров к насыщению, вводят *относительную влажность*.

Относительная влажность — это отношение парциального давления p водяного пара, содержащегося в воздухе при данной температуре, к давлению p_0 насыщенного пара при этой температуре. Относительную влажность обычно выражают в процентах:

$$\varphi = \frac{p}{p_0} \cdot 100\ \%. \tag{1}$$

Если понижать температуру воздуха при неизменном содержании в ней паров воды (т. е. при неизменной абсолютной влажности), то относительная влажность будет возрастать и при некоторой температуре достигнет 100 %, т. е. пар станет насыщенным. Эту температуру, при которой содержащийся в воздухе пар становится насыщенным, называют точкой росы. При охлаждении воздуха до точки росы начинается конденсация паров воды. Это проявляется в образовании тумана и выпадении росы.

Измерение влажности воздуха. Для измерения влажности воздуха существуют приборы гигрометры и психрометры. Действие конденсационного гигрометра основано на измерении точки росы, т. е. на измерении температуры охлаждаемой поверхности в тот момент, когда она запотевает. Зная точку росы и температуру воздуха, можно по специальным таблицам определить парциальное давление паров воды и относительную влажность.

Действие психрометра основано на зависимости интенсивности испарения воды от относительной влажности воздуха. Психрометр состоит из двух термометров, шарик одного из которых сухой, а шарик другого обернут полоской влажной ткани, конец которой погружен в воду. По разности показаний сухого и влажного термометров с помощью специальных таблиц определяют влажность воздуха.

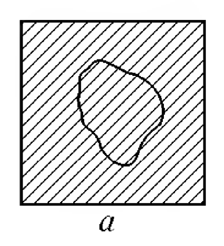
Поверхностное натяжение. В термодинамической системе, состоящей из нескольких фаз, всегда имеются границы раздела. С этими межфазными границами связан ряд свойств системы, к обсуждению которых мы и переходим.

В жидкой фазе на молекулы вещества действуют силы притяжения со стороны окружающих молекул. Когда молекула находится в

глубине жидкости и удалена от поверхности, то эти силы действуют во все стороны и в среднем уравновешиваются. Если же молекула находится вблизи поверхности, то действие на нее других молекул приводит к появлению равнодействующей силы, направленной внутрь жидкости. Для увеличения свободной поверхности жидкости часть молекул должна перейти из объема в поверхностный слой, а для этого необходимо совершить некоторую работу. Поэтому существование поверхности жидкости связано с дополнительной энергией.

Поверхностная энергия — это избыток потенциальной энергии молекул в поверхностном слое по сравнению с энергией молекул внутри фаз (т. е. в объеме вдали от границы), обусловленный отмеченным различием межмолекулярных взаимодействий в граничащих фазах.

Сила поверхностного натяжения. Сила поверхностного натяжения — это сила, которая действует вдоль границы раздела, например вдоль поверхности жидкости, перпендикулярно линии (реальной или мысленной), ограничивающей эту поверхность. В существовании силы поверхностного натяжения можно убедиться с помощью следующих простых опытов.


а

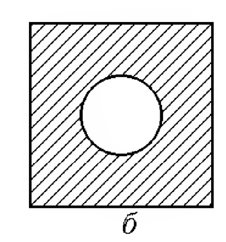

б

Рис. 81. Петля из нитки на мыльной пленке

Наиболее наглядны опыты с использованием мыльной пленки. Положим петлю из нитки на мыльную пленку, полученную при обмакивании проволочного каркаса в мыльном растворе (рис. 81). Пока мыльная пленка внутри петли цела, петля сохраняет любую приданную ей форму (рис. 81*а*). Если же пленку внутри петли проткнуть, то нить принимает форму окружности (рис. 81*б*). Такая форма обеспечивает минимальную площадь оставшейся на каркасе пленки и тем самым минимальное значение ее поверхностной энергии.

Благодаря поверхностному натяжению жидкость в отсутствие внешних сил принимает форму шара, т. е. имеет минимальную при данном объеме поверхность и соответственно минимальную поверхностную энергию. Так бывает в невесомости, например в кабине космического корабля, где невозможно налить воду в стакан, и при свободном падении дождевых капель.

Поверхностная энергия. Количественно *поверхностное натяжение* σ характеризуется отношением модуля F силы поверхностного натяжения, действующей на границу поверхностного слоя, к длине границы l, или, что то же самое, отношением поверхностной энергии $U_{\text{пов}}$ к площади S поверхности:

$$\sigma = \frac{F}{l} = \frac{U_{\text{пов}}}{S}. \qquad (2)$$

В случае плоской поверхности жидкости сила поверхностного натяжения не зависит от того, насколько поверхность «растянута». Строго говоря само понятие «растяжение» для такой поверхности лишено смысла: чтобы увеличивать площадь поверхности, вытягивая в поверхностный слой из объема все новые и новые молекулы, необходимо, в отличие от случая растягивания резиновой пленки, прикладывать постоянную силу, так как поверхностный слой, увеличиваясь по площади, не меняет своих свойств.

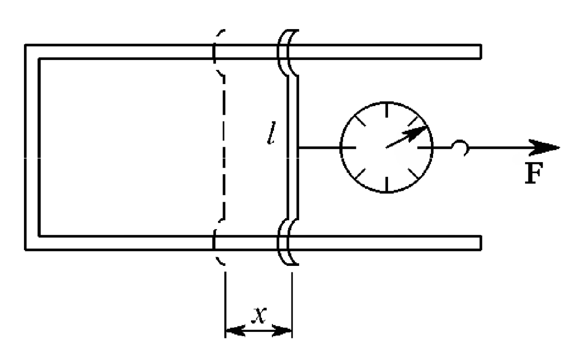


Рис. 82. К вычислению работы при увеличении поверхности пленки

Именно это свойство сил поверхностного натяжения приводит к эквивалентности двух определений величины σ в формуле (2). В самом деле, работа, совершаемая при перемещении ограничивающей мыльную пленку перемычки длиной l на расстояние x (рис. 82) равна $Fx = 2\sigma lx = 2\sigma\Delta S$, где $\Delta S = lx$ — увеличение площади поверхности одной стороны пленки. Эта работа равна приращению поверхностной энергии пленки $\Delta U_{\text{пов}} = 2\sigma\Delta S$.

Смачивание. Вследствие взаимодействия молекул жидкости с молекулами твердых тел возникает *смачивание* — явление, приводящее к искривлению свободной поверхности жидкости у поверхности твердого тела. Если сила взаимодействия между молекулами жидкости и твердого тела больше силы взаимодействия между молекулами самой жидкости, то говорят, что жидкость *смачивает* твердую поверхность. В этом случае угол θ между плоскостью, касательной к поверхности жидкости, и поверхностью твердого тела, называемый краевым углом или углом смачивания, будет острым (рис. 83*а*). В противном случае угол θ будет тупым

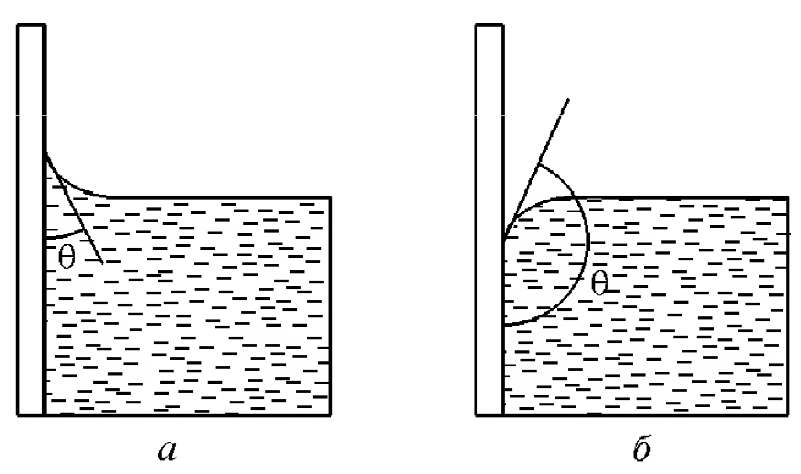


Рис. 83. Краевой угол в случае смачивания (*а*) и несмачивания (*б*)

(рис. 83*б*). При полном смачивании $\theta = 0$, при полном несмачивании $\theta = \pi$. Например, вода смачивает чистое стекло и не смачивает поверхность, покрытую парафином.

Смачиванием и несмачиванием объясняются многие хорошо знакомые нам явления. Подъем керосина по фитилю лампы, возможность вытереться полотенцем — это примеры явлений, целиком обусловленных смачиванием. Напротив, плавание покрытой тонким слоем жира стальной иголки на поверхности воды, тефлоновые кастрюли и сковородки, в которых не подгорает пища, бегающие по воде на длинных тонких ногах жуки-водомеры — во всех этих случаях мы сталкиваемся с несмачиванием.

Капиллярные явления. Подъем или опускание жидкости в узких трубках-капиллярах вследствие явления смачивания называются *капиллярными явлениями* (рис. 84). Высоту h, на которую поднимается жидкость (плотностью ρ), полностью смачивающая

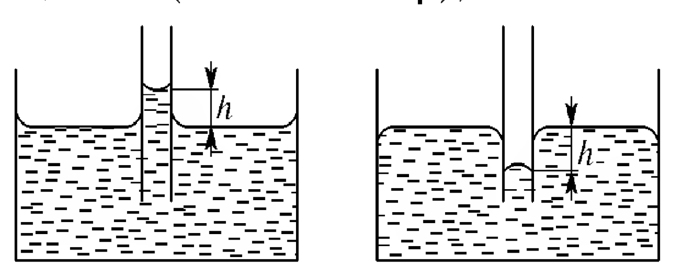


Рис. 84. Жидкость в капиллярах

стенки капилляра радиуса r, можно найти, приравнивая вес столбика поднявшейся жидкости силе поверхностного натяжения, действующей по верхнему периметру столбика:

$$\rho g \pi r^2 h = 2\pi r \sigma,$$

откуда

$$h = \frac{2\sigma}{\rho g r}. \tag{3}$$

Капиллярное давление. Искривление поверхности жидкости, вызываемое силами поверхностного натяжения, ведет к появлению избыточного давления Δp в жидкости, называемого капиллярным. В случае сферической поверхности это давление связано с радиусом ее кривизны r формулой Лапласа:

$$\Delta p = \frac{2\sigma}{r}. \tag{4}$$

Эту формулу легко получить с помощью (3), если учесть, что избыточное гидростатическое давление у основания столба жидкости в капилляре должно компенсироваться изменением давления, обусловленным искривлением свободной поверхности жидкости.

Формулу Лапласа можно получить и иначе, например сравнивая работу $p\Delta V$, совершаемую при выдувании мыльного пузыря, с увеличением поверхностной энергии $\delta \Delta S$ мыльной пленки.

Формулу Лапласа (4) можно применять как к выпуклой, так и к вогнутой поверхности. В случае вогнутой поверхности нужно считать $r < 0$ и соответственно $\Delta p < 0$; для выпуклой, наоборот, $r > 0$ и $\Delta p > 0$.

Задача

На какую высоту поднимется жидкость между двумя параллельными вертикальными стеклами с узким зазором между ними, равным l, если концы стекол погрузить в жидкость, полностью смачивающую стекло?

Решение. При полном смачивании свободная поверхность жидкости в узком зазоре между стеклами примет цилиндрическую форму, причем диаметр $2r$ этого цилиндра равен толщине l зазора. Для определения высоты подъема можно приравнять вес поднявшейся жидкости $\rho glLh$ силе поверхностного натяжения $2\sigma L$, действующей вдоль всей длины $2L$ границы жидкости со стеклом в зазоре:

$$\rho glLh = 2\sigma L.$$

Отсюда

$$h = \frac{2\sigma}{\rho gl}.$$

Обратим внимание на то, что эта высота вдвое меньше высоты подъема в капилляре, диаметр которого равен ширине зазора. Соответственно вдвое меньше и величина капиллярного давления, создаваемого цилиндрической поверхностью по сравнению со сферической.

Капиллярные волны. Силы поверхностного натяжения, как и силы тяжести, играют важную роль в образовании волн на поверхности воды. Роль этих сил различна для волн разной длины: для достаточно коротких волн, когда кривизна поверхности жидкости велика, преобладают силы поверхностного натяжения; в случае длинных волн, наоборот, этими силами можно пренебречь. Поэтому мелкая рябь на воде — это по существу капиллярные волны. Соображения размерности позволяют легко найти характер зависимости скорости v распространения капиллярных волн от плотности жидкости ρ, поверхностного натяжения σ и длины волны λ. Составляя выражение

$$v = C\rho^x\sigma^y\lambda^z$$

и выписывая соответствующее равенство размерностей, находим

$$v = C\sqrt{\frac{\sigma}{\rho\lambda}}.$$

Точная динамическая теория дает для безразмерного коэффициента C значение, равное $\sqrt{2\pi}$.

- Что такое фазы и что такое компоненты термодинамической системы?
- Приведите примеры многофазных однокомпонентных систем и однофазных многокомпонентных систем.
- Какие фазовые превращения вы знаете?
- При каких условиях различные фазы могут находиться в равновесии?

- Что такое насыщенный пар? Объясните, почему его давление не зависит от объема.
- Почему давление насыщенного пара с увеличением температуры растет быстрее, чем давление идеального газа?
- Объясните, почему температура кипения жидкости зависит от внешнего давления. Почему при определенном давлении температура кипения на границе несмешивающихся жидкостей ниже температуры кипения каждой из них?
- Если стакан с водой положить под колокол воздушного насоса, то при быстром откачивании воздуха вода в стакане закипает. Объясните, почему.
- На чем основан принцип действия пузырьковой камеры?
- Что такое критическая температура? Как связано различие понятий «газ» и «пар» с критической температурой?
- Поясните физическую природу теплоты парообразования. Почему она обращается в нуль при критической температуре?
- Чему равна разность показаний сухого и влажного термометров в психрометре при относительной влажности воздуха 100 %?
- Чем обусловлено различие энергии молекул, находящихся в объеме жидкости и на ее поверхности?
- Почему дождевая капля в свободном полете немного сплюснута, а капля, свисающая с водопроводного крана, — вытянута?
- Почему тонкая сплошная вертикальная струя воды, вытекающая из крана, на некотором расстоянии распадается на отдельные капли?
- Почему у пловца под водой волосы свободно развеваются, и сразу слипаются, как только пловец выйдет из воды?
- Какую форму принимает капля воды на горизонтальной деревянной, стеклянной и алюминиевой поверхности?
- Почему капля воды на раскаленной сковороде сохраняет шарообразную форму, несмотря на то, что вода смачивает поверхность металла?
- Где при выводе формулы (3) использовано предположение о том, что жидкость полностью смачивает внутреннюю поверхность капилляра?
- Чем объясняется различие формы поверхности жидкости в узком капилляре и широкой трубке?
- При подъеме жидкости в капилляре до высоты h силы поверхностного натяжения совершают работу mgh, где m — масса столба жидкости (докажите это). Но потенциальная энергия этого столба равна только $mgh/2$. Как это согласуется с законом сохранения энергии?

§ 26. Реальные газы. Уравнение Ван-дер-Ваальса

Поведение реальных газов в той или иной степени отличается от поведения идеального газа, описываемого уравнением Менделеева—Клапейрона. Отклонения зависят не только от того, с каким газом — кислородом, азотом и т. д. — мы имеем дело, но и от тех условий, в которых находится газ. Чем более разрежен газ и чем

выше его температура, тем менее заметны эти отклонения. Поэтому применимость модели идеального газа к какому-либо реальному газу определяется не только свойствами самого газа, такими, как размеры и масса его молекул и взаимодействия между ними, но и условиями, в которых находится газ.

О моделях реального газа. Взаимодействие между молекулами того или иного реального газа имеет настолько сложный характер, что невозможно получить уравнение состояния, которое количественно правильно описывало бы поведение реального газа во всей области возможных изменений его температуры и плотности. Можно, однако, написать приближенное уравнение состояния реального газа, учитывающее основные качественные особенности межмолекулярного взаимодействия.

В модели идеального газа не учитывались ни взаимодействие молекул на расстоянии, ни конечный размер самих молекул. Напомним, что силы взаимодействия между молекулами имеют электромагнитную природу, хотя в целом молекулы электрически нейтральны. На больших расстояниях преобладают силы притяжения между молекулами, а на малых — силы отталкивания. Силы притяжения между молекулами быстро убывают с увеличением расстояния между ними, и поэтому при малой плотности газ ведет себя как идеальный. Но тем не менее силы притяжения существенны, так как именно они приводят к конденсации газа в жидкость. Отталкивание молекул на малых расстояниях настолько быстро возрастает при их сближении, что с хорошей точностью молекулы можно считать твердыми взаимно не проникающими телами и говорить об их собственном объеме и эффективном диаметре. Этот объем ставит предел возможному сжатию газа.

Модель твердых шаров. Приведенные соображения позволяют предложить простейшую микроскопическую модель реального газа — модель твердых шаров. В этой модели молекулы рассматриваются как твердые шары радиуса r_0, взаимодействующие только при соприкосновении друг с другом. Кроме радиуса r_0, система характеризуется массой m молекул (шаров), их концентрацией n и средней кинетической энергией теплового движения $(3/2)kT$.

Чтобы разобраться в свойствах такой модели, найдем соответствующий ей безразмерный параметр γ, составленный из r_0, m, n и kT. Поскольку время входит только в kT, то kT в безразмерный параметр γ входить не может. Но тогда не может входить в него и масса m. А из оставшихся величин r_0 и n можно составить только один независимый безразмерный параметр $\gamma = r_0^3 n$, физический смысл которого — отношение собственного объема шара к среднему объему, приходящемуся на одну молекулу в газе. Предельному случаю идеально-

го газа соответствует сильное разрежение, т. е. $\gamma \ll 1$, когда энергия взаимодействия молекул должна стать пренебрежимо малой по сравнению с энергией теплового движения kT.

В сильно сжатом газе отношение средней энергии взаимодействия молекул к их кинетической энергии должно зависеть от температуры, поскольку с уменьшением температуры отклонение реального газа от идеального становится все более и более заметным. Однако в модели твердых шаров это отклонение определяется только параметром γ и, следовательно, не зависит от температуры. Поэтому последовательная микроскопическая модель реального газа должна учитывать взаимодействие молекул друг с другом и на расстоянии, а не только при непосредственном соприкосновении.

Модель Ван-дер-Ваальса. Приступая к получению приближенного уравнения состояния реального газа, будем считать, что взаимодействие молекул приводит лишь к небольшим поправкам в уравнении состояния идеального газа. При достаточно высоких температурах и малых плотностях газа искомое уравнение должно приводить к тем же результатам, что и уравнение Менделеева—Клапейрона.

Учет размеров молекул. Прежде всего учтем конечные собственные размеры молекул. Фактически это приводит к тому, что предоставленный молекулам газа для движения объем будет меньше объема сосуда V. Поэтому в уравнении состояния одного моля идеального газа заменим объем V на $V - b$, где b — характерная для данного газа положительная постоянная, учитывающая занимаемый молекулами объем:

$$p(V - b) = RT. \tag{1}$$

Из этого уравнения видно, что объем газа V не может быть сделан меньше, чем b, так как при $V \to b$ давление газа p неограниченно возрастает. Разумеется, не следует ожидать, что при значениях V, близких к b, уравнение (1) будет правильно описывать поведение газа, ибо по своему смыслу постоянная b является малой поправкой: $b \ll V$.

Учет взаимодействия между молекулами. Учтем теперь проявляющееся на больших расстояниях притяжение между молекулами. Притяжение должно приводить к уменьшению оказываемого газом давления на стенки сосуда, так как на каждую находящуюся вблизи стенки молекулу будет действовать со стороны остальных молекул газа сила, направленная внутрь сосуда. Поэтому давление на стенки p будет меньше значения, даваемого выражением (1), на некоторую величину p':

$$p = \frac{RT}{V - b} - p'. \tag{2}$$

От чего может зависеть величина p'? Можно ожидать, что в грубом приближении сила, действующая на каждую молекулу со

стороны всех остальных, будет пропорциональна числу окружающих молекул, т. е. плотности газа. Следовательно, и поправка к передаваемому *отдельной* молекулой импульсу при ударе о стенку пропорциональна плотности газа. Поправка же к передаваемому при ударах о стенку всеми молекулами импульсу (т. е. к давлению) будет пропорциональна *квадрату* плотности газа или, что то же самое, обратно пропорциональна квадрату объема. Поэтому выражение для p' можно записать в виде

$$p' = \frac{a}{V^2}, \tag{3}$$

где a — характерная для данного газа положительная постоянная. Поскольку p', как и b, является малой поправкой, должно выполняться неравенство $a/V^2 \ll p$. Подставляя p' из формулы (3) в уравнение (2), получим

$$\left(p + \frac{a}{V^2}\right)(V - b) = RT. \tag{4}$$

Уравнение (4) называется *уравнением Ван-дер-Ваальса.* Оно приближенно учитывает особенности поведения реального газа, обусловленные межмолекулярным взаимодействием и собственными размерами молекул. Постоянные Ван-дер-Ваальса a и b определяются опытным путем: их значения для каждого газа выбираются таким образом, чтобы уравнение (4) наилучшим образом описывало поведение данного газа.

Феноменологический характер уравнения Ван-дер-Ваальса. Не следует рассматривать приведенные выше рассуждения как строгий вывод уравнения состояния реального газа. Они представляют собой пример феноменологического подхода, при котором качественный вид закономерности устанавливается с помощью наводящих соображений, а количественные характеристики — в данном случае постоянные a и b — находятся из сравнения с экспериментом. Как и всякое феноменологическое соотношение, уравнение Ван-дер-Ваальса в некоторой области — при достаточно высоких температурах и малых плотностях — дает правильное количественное описание свойств реального газа, тогда как во всей области изменения параметров оно дает только качественную картину поведения газа.

Кроме уравнения Ван-дер-Ваальса было предложено много других эмпирических уравнений состояния реальных газов. Некоторые из них дают лучшее согласие с опытом за счет большего числа входящих в них феноменологических постоянных. Однако при качественном исследовании поведения реальных газов удобно использовать именно уравнение Ван-дер-Ваальса благодаря его простоте и ясному физическому смыслу.

Изотермы Ван-дер-Ваальса. Для исследования поведения газа, подчиняющегося уравнению Ван-дер-Ваальса, рассмотрим определяемые этим уравнением изотермы, т. е. кривые зависимости p от V при заданных значениях температуры T. С этой целью перепишем уравнение (4) в виде

$$V^3 - \left(b + \frac{RT}{p}\right)V^2 + \frac{a}{p}\,V - \frac{ab}{p} = 0. \qquad (5)$$

При фиксированных значениях p и T это уравнение третьей степени относительно V. Уравнение третьей степени имеет либо один, либо три вещественных корня. Поэтому при данных значениях давления и температуры уравнение (5) дает либо одно, либо три значения объема. Это значит, что на p—V-диаграмме изотерма пересекает горизонтальную прямую $p = \text{const}$ либо в одной точке, либо в трех точках.

На рис. 85 приведены изотермы, соответствующие различным значениям температуры $T_1 > T_{\text{кр}} > T_2 > T_3$. Изотерма, соответствующая достаточно высокой температуре T_1, мало отличается от изотермы идеального газа, и одно из самых существенных отличий заключается

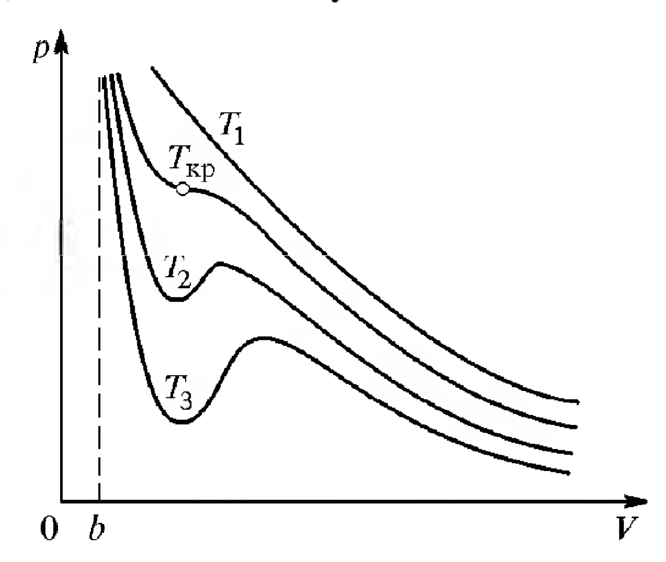


Рис. 85. Изотермы газа Ван-дер-Ваальса на p—V-диаграмме

в том, что значение объема V ни при каких давлениях не может быть меньше b. При достаточно низких температурах T_2 и T_3 изотермы содержат волнообразный участок, так что прямая $p = \text{const}$ трижды пересекает изотерму. Существует некоторая температура $T_{\text{кр}}$, такая, что соответствующая ей изотерма отделяет все монотонные изотермы, лежащие выше нее, от «горбатых», лежащих ниже. Такая температура и соответствующая ей изотерма носят название *критических*.

Экспериментальные изотермы Эндрюса. Для того чтобы выяснить смысл этой на первый взгляд весьма странной зависимости, следует обратиться к опыту. На рис. 86 показаны полученные эксперимен-

тально изотермы углекислого газа CO_2. При высоких температурах эти изотермы очень похожи на изотермы идеального газа. При низ-

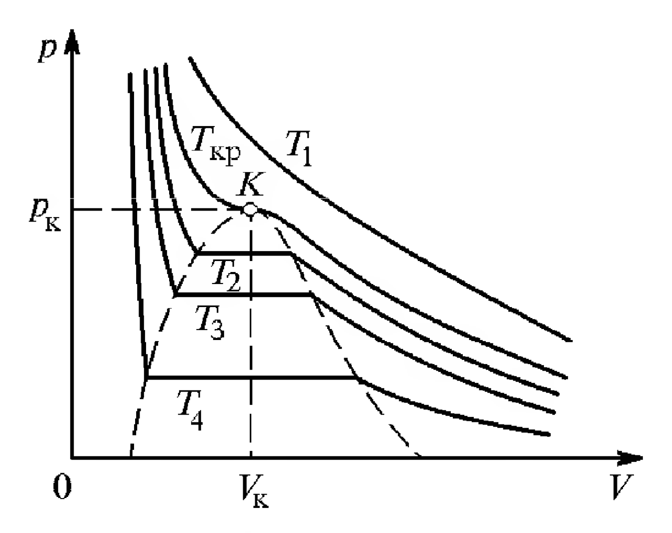


Рис. 86. Экспериментальные изотермы углекислого газа

ких температурах характер изотерм существенно другой: у них имеется горизонтальный участок, на котором давление остается постоянным, несмотря на изменение объема.

Рассмотрим, каким физическим процессам соответствуют различные участки изотерм. При изотермическом сжатии одного моля углекислоты на участке *AB* (рис. 87) она ведет себя подобно идеальному га-

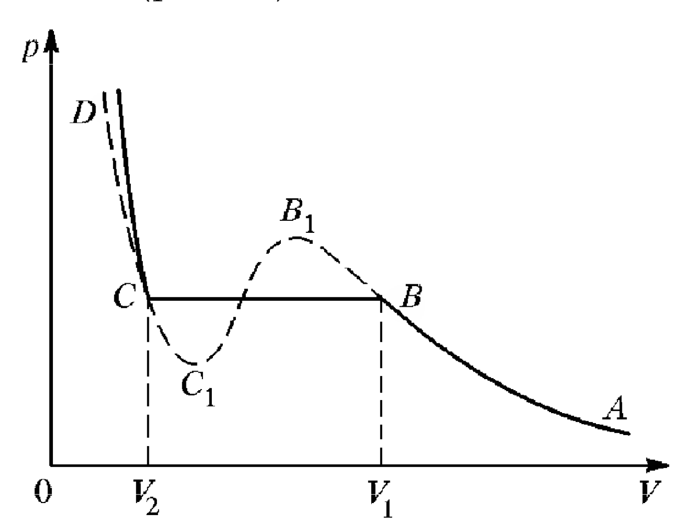


Рис. 87. Сравнение экспериментальной изотермы с изотермой Ван-дер-Ваальса

зу: давление возрастает при уменьшении объема. При дальнейшем сжатии давление в сосуде остается неизменным, но в сосуде появляется жидкая углекислота. Так продолжается на всем участке *BC*, пока в точке *C* весь газ не сконденсируется в жидкость. Участок *BC* соответствует равновесию между жидкой углекислотой и ее насыщенным паром. Дальнейшие попытки сжатия жидкости сопровождаются резким увеличением давления при самом незначительном уменьшении объема.

Наложим на график экспериментально полученной изотермы теоретическую кривую Ван-дер-Ваальса для той же самой температуры. Параметры a и b можно выбрать так, чтобы на участке AB, соответствующем газообразному состоянию, обе кривые практически совпадали. Это значит, что уравнение Ван-дер-Ваальса вполне удовлетворительно описывает свойства паров, далеких от насыщения. Крутой ход теоретической изотермы на участке CD качественно правильно передает малую сжимаемость жидкости, хотя экспериментальная изотерма на этом участке идет еще круче. Но волнообразный участок теоретической изотермы совершенно не похож на соответствующее ему плато экспериментальной кривой.

Метастабильные состояния. Однако оказывается, что ряд состояний, попадающих в область «горбов», может быть экспериментально осуществлен при соблюдении определенных условий. Так, в свободном от пыли и заряженных частиц пространстве удается при постепенном сжатии получить пары при давлении, большем давления насыщенных паров при данной температуре. Такие пары называются *перенасыщенными*, и их состояния достаточно хорошо описываются участком BB_1 теоретической изотермы. Эти состояния при появлении центров конденсации — ионов, пылинок, мельчайших капелек — становятся неустойчивыми, и пересыщенный пар мгновенно конденсируется в туман.

Пересыщенный пар можно получить не только при изотермическом сжатии, но и при охлаждении насыщенного пара. Поэтому его иногда называют также переохлажденным. Такой способ получения перенасыщенного пара используется в камере Вильсона, нашедшей широкое применение на заре ядерной физики для регистрации заряженных частиц.

Жидкая углекислота при постепенном изотермическом расширении может сохраниться в жидком состоянии и при давлении, меньшем давления насыщенных паров при данной температуре. Такие состояния, как уже отмечалось, называются перегретой жидкостью, и им соответствует участок CC_1 теоретической изотермы Ван-дер-Ваальса. Перегретая жидкость может быть получена также нагреванием жидкости при постоянном давлении, если принять меры к тому, чтобы жидкость и стенки сосуда не содержали растворенных газов, которые при нагревании могут выделяться с образованием пузырьков и становиться зародышами новой фазы. При появлении центров кипения перегретая жидкость мгновенно закипает.

Таким образом, участки AB_1 и DC_1 теоретической изотермы соответствует состояниям вещества, которые, хотя и неустойчивы, но могут быть осуществлены экспериментально. Такие состояния называются метастабильными.

Абсолютно неустойчивые состояния. Состояния вещества на участке C_1B_1 (см. рис. 87) являются абсолютно неустойчивыми и вообще не могут осуществляться в природе, так как им соответствует уменьшение давления при сжатии. Представим на минуту, что такие состояния возможны. Пусть небольшая часть вещества случайно сжалась в результате флуктуации плотности. Тогда давление в этом месте уменьшится и станет меньше давления окружающей среды. Это приведет к дальнейшему сжатию выделенного вещества и т. д. — система самопроизвольно выходит из такого неустойчивого состояния.

Если неустойчивость состояний, соответствующих участкам CC_1 и BB_1, обусловливалась наличием центров конденсации и кипения и потому была в принципе устранимой, то неустойчивость состояний на участке C_1B_1 связана с неизбежными тепловыми флуктуациями и принципиально неустранима. Поэтому такие состояния теоретической изотермы никогда не осуществляются.

Наличие неосуществимого участка на изотермах Ван-дер-Ваальса при температурах ниже критической означает, что при постепенном изменении объема вещество не может все время оставаться однородным: в некоторый момент должно произойти расслоение вещества на две фазы — жидкую и газообразную.

В связи с этим интересно отметить, что эмпирическое уравнение Ван-дер-Ваальса, полученное с целью введения малых поправок к уравнению состояния идеального газа, фактически оказалось эффективным в гораздо более широкой области. В самом деле, оно указало на существование критической температуры и на необходимость расслоения вещества на фазы при температурах ниже критической. Оно также отразило возможность существования состояний пересыщенного пара и перегретой жидкости, качественно описало малую сжимаемость жидкостей и т. п.

Критическое состояние вещества. Проследим, как изменяется по мере повышения температуры область сосуществования двух фаз на p—V-диаграмме. Прямолинейный участок экспериментальной изотермы уменьшается с ростом температуры и при критической температуре стягивается в одну точку K на рис. 86. Если соединить между собой точки начала горизонтальных участков на всех изотермах, а также точки концов этих участков, то получим некоторую кривую MKN (рис. 88). Эта кривая представляет собой границу, отделяющую состояния, в кото-

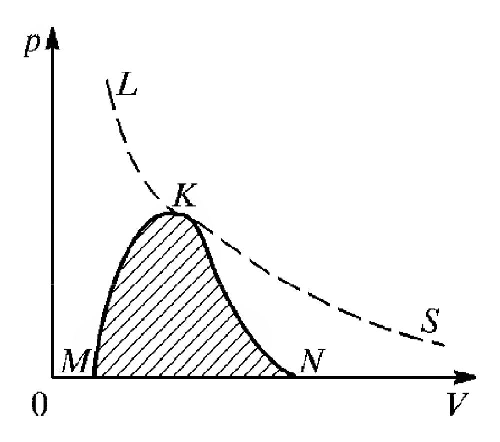


Рис. 88. Границы равновесия различных фаз вещесва на p—V-диаграмме

рых вещество существует в двух находящихся в равновесии фазах, от однофазных состояний. Соответствующие точке K давление и объем получили название критических. Состояние вещества в точке K называется *критическим состоянием*.

В критической точке сливаются в одну три точки, в которых горизонтальный участок экспериментальной изотермы пересекает изотерму Ван-дер-Ваальса.

В критическом состоянии исчезает различие между жидкостью и газом, нет и никакой границы раздела между ними. При этом теряют смысл такие понятия, как поверхностное натяжение, теплота парообразования.

Сжимаемость вещества при приближении к критическому состоянию неограниченно возрастает: график изотермы в окрестности этой точки идет горизонтально и, следовательно, малое изменение объема вещества не сопровождается изменением давления. Это приводит к существованию больших флуктуаций плотности в критическом состоянии, которые проявляются, например, в явлении критической опалесценции — сильном рассеянии света на случайных неоднородностях.

Границы равновесия разных фаз вещества между собой показаны на диаграмме состояний (рис. 88). Граница KS является чисто условной и соответствует историческому делению газообразного состояния на пар и «истинный газ» в зависимости от того, находится ли вещество при температуре ниже или выше критической. В отличие от границы MKN на этой диаграмме, при переходе через границу LKS никаких качественных изменений в веществе не происходит.

- В каких явлениях обнаруживаются отклонения поведения реальных газов от предсказаний модели идеального газа?
- Почему модель твердых шаров не в состоянии объяснить факт усиления отклонения свойств реального газа от идеального при понижении температуры?
- Какой физический смысл имеют параметры a и b, фигурирующие в уравнении Ван-дер-Ваальса? Какова их размерность? Каким образом определяются их значения для реальных газов?
- В чем заключается феноменологический характер уравнения Ван-дер-Ваальса?
- С чем связан качественно различный вид изотерм Ван-дер-Ваальса при разных температурах?
- Опишите физические процессы при изменении объема системы, соответствующие разным участкам экспериментальной изотермы.
- Как практически можно осуществить метастабильные состояния, соответствующие участкам CC_1 и BB_1 теоретической изотермы Ван-дер-Ваальса?
- Почему участок B_1C_1 теоретической изотермы принципиально неосуществим на опыте?

- Поясните, каким образом уравнение Ван-дер-Ваальса указывает на необходимость расслоения вещества на фазы при температурах ниже критической.
- Что такое критическое состояние? Какими свойствами обладает вещество в критическом состоянии?
- Поясните, почему при переходе через границу *LKS* на рис. 88, разделяющую газ и пар, никаких качественных изменений в состоянии вещества не происходит.

§ 27. Фазовые переходы

Переход вещества из одной фазы в другую при заданном давлении происходит всегда при строго определенной температуре. Например, если нагревать лед при нормальном атмосферном давлении, то он начнет плавиться при достижении температуры 0 °C, и эта температура, несмотря на подвод теплоты, будет оставаться неизменной до тех пор, пока весь лед не растает и не превратится в воду. В процессе плавления лед и вода существуют одновременно, находясь в тепловом равновесии друг с другом. Если прекратить подвод теплоты, то лед и вода будут сосуществовать при этой температуре неограниченно долго.

Диаграмма состояний. Если изменить давление, то изменится и температура фазового перехода. Другими словами, сосуществование двух фаз в равновесии возможно лишь при строго определенном соотношении между давлением и температурой. Если изобразить эту зависимость в виде кривой в системе координат p—T, то получится так называемая *диаграмма состояний.*

Кривая равновесия пар–жидкость. Рассмотрим диаграмму состояний для фазового перехода между жидкостью и ее паром. Вспомним, что на p—V-диаграмме (рис. 86) равновесию между жидкостью и паром соответствуют горизонтальные участки экспериментальных изотерм. Горизонтальный ход изотермы на этих участках соответствует тому, что при данной температуре независимо от объема равновесие между жидкостью и паром возможно только при определенном давлении. Поэтому для получения кривой равновесия жидкости и пара на p—T-диаграмме нужно на оси абсцисс отложить значения температуры для каждой изотермы, а на оси ординат давление, соответствующее ее горизонтальному участку. В результате получится кривая, изображенная на рис. 89: давление, при котором

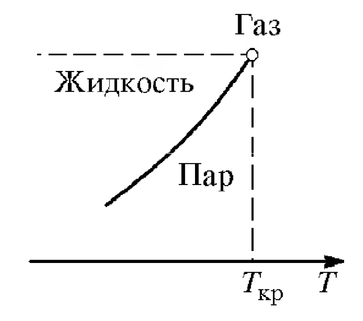


Рис. 89. Кривая равновесия жидкости и пара на p—T-диаграмме

жидкость и пар могут находиться в равновесии, возрастает с увеличением температуры.

Кривая сосуществования заканчивается в критической точке, так как из p—V-диаграмм видно, что по мере повышения температуры горизонтальный участок изотерм становится все меньше и исчезает при критической температуре — при температурах выше критической жидкость существовать не может. По мере приближения к критической температуре плотность насыщенного пара возрастает; при $T = T_{кр}$ она становится равной плотности жидкости и пар вообще становится неотличимым от жидкости.

Кривая равновесия жидкости и пара на p—T-диаграмме дает зависимость давления насыщенных паров от температуры. В то же самое время эта кривая дает зависимость температуры кипения жидкости от давления, так как кипение происходит при такой температуре, когда давление насыщенного пара в пузырьках становится равным давлению в жидкости. Поэтому иногда она называется кривой кипения.

Кривая равновесия жидкости и пара представляет собой некоторую границу на p—T-диаграмме. Так как при заданном давлении более высоким температурам соответствует пар, а более низким температурам — жидкость и, наоборот, при заданной температуре более низким давлениям соответствует пар, а более высоким давлениям — жидкость, то область справа и снизу от кривой равновесия соответствует газообразной, а область слева и сверху от нее — жидкой фазе.

Переход пар–жидкость. Существование критической точки на кривой равновесия жидкости и пара отчетливо демонстрирует отсутствие принципиальной разницы между жидким и газообразным состояниями вещества. Поясним это следующим примером. Возьмем пар в некотором состоянии A (рис. 90) и переведем его в жидкость в состоянии, изображенном точкой B. Такое превращение можно осуществить разными путями. Можно, например, охлаждать газ при постоянном давлении p_A до тех пор, пока температура не станет равной T_B, а затем поднять давление при постоянной температуре до значения p_B (путь *1* на рис. 90). В процессе изобарического охлаждения путь *1* пересекает кривую равновесия, и в этой точке система расслаивается на две фазы, так как происходит конденсация пара в жидкость.

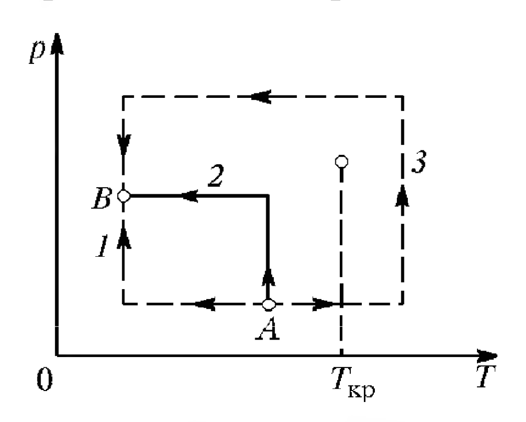


Рис. 90. Переход из газообразного состояния A в жидкое B по разным путям на p—T-диаграмме

Можно перевести систему из состояния A в состояние B по пути *2*, сначала изотермически сжимая газ до давления p_B, а затем изо-

барически охлаждая до температуры T_B. В этом случае конденсация пара и расслоение на две фазы произойдет в процессе изотермического сжатия.

Но переход между теми же состояниями A и B можно провести и таким способом, при котором нигде не будет происходить скачкообразного изменения состояния и вещество все время будет оставаться однородным. Этого можно добиться, осуществляя процесс в обход кривой сосуществования пара и жидкости, например по пути *3* на рис. 90. Осуществляя переход по этому пути, мы нигде не увидели бы конденсации и, следовательно, не смогли бы сказать, что вещество перестало быть газом и стало жидкостью. И тем не менее и в этом случае мы попадем в то же самое конечное состояние, в котором рассматриваемую систему мы считаем жидкостью.

В обычных условиях газы и жидкости настолько сильно отличаются друг от друга по плотности, что не представляет никакого труда различить их. Но, как мы видели, различие между этими состояниями вещества в действительности не принципиальное, а скорее количественное. Различие в плотности сказывается на интенсивности взаимодействия молекул и на характере теплового движения. Но и газ при температурах выше критической можно сжать до такой степени, что его плотность станет больше характерного значения плотности жидкости.

Рис. 91. Запаянная ампула с жидкостью и ее насыщенным паром

Изохорический переход пар–жидкость. Рассмотрим теперь процесс перехода вещества из одной фазы в другую, происходящий при неизменном объеме. Будем нагревать запаянную ампулу, содержащую некоторое количество жидкости и находящийся над ее поверхностью насыщенный пар (рис. 91). Если внутренний объем ампулы V_1 превышает критический объем $V_{кр}$, соответствующий находящемуся в ампуле количеству вещества (рис. 92), то прямая, соответствующая изохорическому нагреванию, пройдет справа от критической точки и по мере нагревания количество жидкости в ампуле будет уменьшаться, пока все вещество не обратится в пар в точке C_1. При этом граница жидкости опускается и исчезает у нижнего конца ампулы. Если объем ампулы V_2 мень-

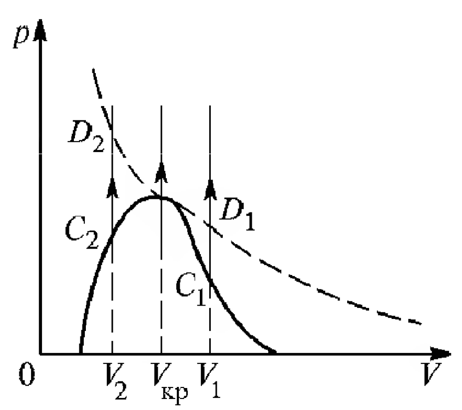


Рис. 92. Поведение границы раздела в зависимости от объема ампулы

ше критического объема $V_{кр}$, то при нагревании пар будет конденсироваться, пока все вещество в ампуле не превратится в жидкость. Это произойдет в точке C_2. Граница раздела жидкости и пара поднимается и исчезает у верхнего конца ампулы.

Наблюдение критического состояния. Если, наконец, объем ампулы равен критическому, то граница исчезнет где-то посередине ампулы, и произойдет это, когда температура станет равной критической. Если объем ампулы лишь немного отличается от критического, например меньше критического, то при нагревании граница жидкость — пар в ампуле перемещается вверх, но исчезает все-таки раньше, чем доходит до верхнего края ампулы.

При объемах ампулы V_1 и V_2, заметно отличающихся от критического объема $V_{кр}$, исчезновение границы раздела в точках C_1 и C_2 происходит при температурах ниже критической. При дальнейшем нагревании ампулы никаких видимых изменений вещества не происходит, в том числе и в точках D_1 и D_2 при температуре, равной критической.

Кривая равновесия жидкой и твердой фаз. Аналогичным образом на p—T-диаграмме состояний можно рассмотреть кривую равновесия жидкой и твердой фаз. Эта кривая выражает зависимость температуры плавления от давления и называется *кривой плавления*. По сравнению с кривой равновесия жидкости и пара она имеет две характерные особенности. Во-первых, на кривой плавления отсутствует критическая точка. Это связано с тем, что истинно твердые, т. е. кристаллические, тела обладают упорядоченной молекулярной структурой и принципиально отличаются от жидкостей и газов своей анизотропией. Переход между жидкостью и кристаллом не может быть произведен непрерывным образом, как это можно сделать для жидкости и газа в обход критической точки. Всегда можно точно указать, к какой фазе — жидкой или кристаллической — относится то или иное состояние вещества.

Во-вторых, наклон кривой плавления на p—T-диаграмме может быть разным, в отличие от кривой кипения, где давление всегда возрастает с увеличением температуры. Если при плавлении объем вещества увеличивается, то кривая плавления наклонена в ту же сторону, что и кривая кипения. Напомним, что при испарении объем всегда возрастает. Наклон кривой плавления будет противоположным, если при плавлении объем уменьшается, как это происходит, например, при таянии льда и плавлении чугуна.

Тройная точка. Две фазы вещества могут находиться в равновесии друг с другом вдоль кривых сосуществования на p—T-диаграмме состояний. Три фазы одного и того же вещества могут одновременно находиться в равновесии друг с другом лишь в одной определенной точке на диаграмме p—T. В этой точке сходятся вместе кривые плавления и кипения. Точки равновесия трех фаз называются тройными точками. Например, у воды одновременное существование в равновесии льда, воды и пара возможно только при давлении 4,62 мм рт. ст. и температуре +0,01 °С.

Тройная точка воды очень удобна в качестве стандартной точки температурной шкалы, так как ее воспроизведение, в отличие от точки кипения или плавления, не требует специального поддержания определенного давления.

Схематическая диаграмма состояний для вещества, имеющего три фазы — твердую, жидкую и газообразную, показана на рис. 93. Наклон кривой плавления AB изображен для случая, когда плавление сопровождается уменьшением объема, как у льда. В отличие от кривой равновесия пара и жидкости AK, оканчивающейся при критической температуре в точке K, кривая плавления AB продолжается неограниченно.

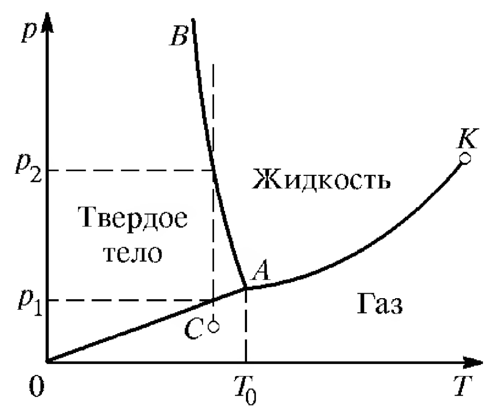


Рис. 93. p—T-диаграмма равновесных состояний вещества

Кривая равновесия твердого тела с газом на p—T-диаграмме уходит в начало координат. Действительно, по законам классической механики при стремлении температуры к абсолютному нулю тепловое движение прекращается. Неподвижные атомы занимают такое расположение, при котором потенциальная энергия взаимодействия минимальна. Это расположение представляет собой регулярную пространственную решетку. Поэтому при абсолютном нуле температуры любое вещество с точки зрения классической механики должно быть кристаллическим.

Квантовые жидкости. Существует только одно исключение из этого правила: гелий остается жидким при всех температурах вплоть до абсолютного нуля, если давление не превышает 25 атм. Это исключение объясняется квантовыми эффектами. Вследствие малой массы атома гелия его дебройлевская длина волны $h/(2mkT)^{1/2}$ при низких температурах (порядка нескольких кельвинов) оказывается сравнимой со средним расстоянием между атомами. В результате присущие атомам гелия волновые свойства и определяют поведение гелия при низких температурах. У всех других одноатомных веществ деб-

ройлевская длина волны настолько мала по сравнению с межатомным расстоянием, что их атомы можно рассматривать как классические частицы вплоть до очень низких температур, при которых происходит кристаллизация.

Говорят, что при низких температурах гелий представляет собой *квантовую жидкость*. Свойства квантовых жидкостей, образованных различными изотопами $^3_2$He и $^4_2$He, оказываются совершенно различными. Это объясняется различием статистических свойств атомов $^3_2$He и $^4_2$He, подчиняющихся квантовым закономерностям.

Испарение твердого тела. Диаграмма состояний позволяет сразу ответить на вопрос, что произойдет с веществом при его нагревании или сжатии. Если, например, вещество в газообразном состоянии, изображаемом точкой C на рис. 93, подвергнуть изотермическому сжатию, то происходящий с ним процесс изобразится вертикальной штриховой линией. Видно, что при давлении p_1 газ затвердеет, а образовавшийся кристалл при давлении p_2 расплавится.

Из диаграммы существования фаз видно, что кристалл при изобарическом нагревании не обязательно должен проходить через стадию жидкого состояния для того, чтобы превратиться в газ. Если давление выше тройной точки, то при нагревании кристалл действительно сначала расплавится, а получившаяся жидкость затем испарится. Но при давлении ниже тройной точки кристалл при нагревании сразу превращается в пар. Такой переход называется *сублимацией* или возгонкой. Именно так ведет себя твердая углекислота при нормальном атмосферном давлении, так как ее тройной точке соответствует давление 5,1 атм.

Правило фаз Гиббса. Итак, в однокомпонентной системе, т. е. в веществе определенного химического состава, все три фазы могут находиться в равновесии. В многокомпонентной системе в равновесии может одновременно находиться большее число фаз. Как было установлено Гиббсом, в системе, состоящей из N компонент, может находиться одновременно в равновесии не больше, чем $N+2$ фаз. Это максимальное число фаз сосуществует в единственном состоянии. Меньшее число фаз могут сосуществовать вдоль кривых равновесия, т. е. при разных давлениях и температурах.

- Как с помощью семейства экспериментальных изотерм получить кривую равновесия пар-жидкость на $p-T$-диаграмме?
- Какие экспериментальные факты свидетельствуют об отсутствии принципиальной разницы между жидким и газообразным состояниями вещества, несмотря на существование четкой границы между ними — поверхности жидкости?

- Каким образом можно перевести жидкость в газ, минуя расслоение вещества на фазы?
- Изобразите качественную картину кривой равновесия жидкой и газообразной фаз для многокомпонентной системы, например для воздуха.
- Как ведет себя граница раздела жидкости и газа в запаянной ампуле при переходе через критическую температуру, если объем ампулы больше или меньше критического объема вещества, находящегося внутри ампулы?
- Что происходит с поверхностным натяжением по мере приближения температуры к критической?
- Почему гелий при нормальном атмосферном давлении остается жидким вплоть до самых низких температур?
- Почему три фазы некоторого количества вещества могут находиться в равновесии между собой только в тройной точке, т. е. при строго определенных значениях температуры, давления и объема?
- При каких условиях кристалл при повышении температуры превращается в газ, минуя жидкую фазу?

VI. АТОМЫ И ИЗЛУЧЕНИЕ

§ 28. Излучение света атомами

В предыдущей главе основы статистической механики применялись для описания наблюдаемых свойств таких макроскопических систем многих частиц, где каждая из частиц подчинялась законам классической механики. В данной главе будут рассматриваться макроскопические системы, состоящие из частиц, подчиняющихся квантовым закономерностям.

Квантовая механика позволяет только на основании известного заряда атомного ядра, а тем самым и числа электронов, найти стационарные состояния электронов в свободном атоме и рассчитать многие свойства атомов, в том числе их оптические спектры. Кроме того, на основе квантовой механики можно рассчитать изменение стационарных состояний атома под действием внешних электрических и магнитных полей. Экспериментально эти изменения проявляются в спектрах испускания и поглощения света атомами. Изменения спектра под действием внешнего магнитного поля (явление Зеемана) и под действием электрического поля (явление Штарка) являются эффективными средствами изучения структуры вещества.

Размер атома и длина волны. Излучаемый атомом свет имеет длину волны, составляющую несколько тысяч ангстрем, что на три порядка превосходит размер атома. Покажем, что большая длина волны излучения является прямым следствием малости постоянной тонкой структуры $\alpha = e^2/\hbar c$, характеризующей силу взаимодействия электрона с электромагнитным полем.

Расстояние между уровнями энергии оптического электрона, определяющее согласно формуле $E_2 - E_1 = \hbar\omega$ частоту излучаемого атомом света, по порядку величины равно энергии связи электрона в атоме водорода $me^4/\hbar^2$. Поэтому длина волны излучения λ равна

$$\lambda = \frac{2\pi c}{\omega} \approx 2\pi \frac{c\hbar^3}{me^4}. \tag{1}$$

Выделим в этом выражении боровский радиус $a_0 = \hbar^2/me^2$, характеризующий размер излучающего атома. Тогда остающийся

множитель $\hbar c/e^2$ равен обратной величине постоянной тонкой структуры:

$$\lambda = \frac{2\pi}{\alpha} a_0 \approx 1000 a_0. \qquad (2)$$

Большая длина волны излучения по сравнению с размером атома имеет, как мы увидим, важные физические следствия, заключающиеся в том, что спектральные линии излучения имеют малую ширину, а время жизни атома в возбужденном состоянии велико по сравнению с периодом колебаний в излучаемой волне.

Стационарные состояния атомов, определяемые квантовой механикой, являются дискретными. Им соответствуют строго определенные значения энергии. Находящийся в возбужденном стационарном состоянии атом в отсутствие внешних воздействий, согласно квантовой механике, должен оставаться в этом состоянии сколь угодно долго. Однако опыт показывает, что это не так: спустя некоторое время возбужденный атом самопроизвольно переходит в основное состояние, испуская при этом квант света. Такое излучение называется *спонтанным*.

Спонтанное излучение. Рассмотрим спонтанное излучение атомов подробнее. Покажем, что при спонтанном переходе в основное состояние атом испускает цуг волн конечной протяженности. Это в свою очередь означает, что излучаемый им свет не является строго монохроматическим, а распределен в некотором частотном интервале. Другими словами, спектральные линии, соответствующие спонтанному излучению атома, имеют некоторую конечную ширину. Это так называемая *естественная ширина* спектральных линий.

Для простоты будем рассматривать случай, когда спонтанное излучение происходит при переходах электрона между состояниями с большими квантовыми числами. Тогда на основании принципа соответствия процесс излучения света электроном можно рассматривать классически. Как было показано, предсказываемый квантовой теорией спектр излучения атома водорода при больших квантовых числах совпадает со спектром, предсказываемым классической теорией излучения: обращающийся по орбите электрон излучает свет, частота которого равна частоте обращения электрона.

Классическая модель спонтанного излучения. Излучая электромагнитные волны, атом теряет энергию. Радиус орбиты электрона при этом постепенно уменьшается, что на квантовом языке соответствует переходу из состояния с более высокой энергией в более низкое квантовое состояние. Оценим время, в течение которого происходит этот переход.

Движение электрона по круговой орбите можно представить как суперпозицию двух гармонических взаимно перпендикулярных колебаний, сдвинутых друг относительно друга по фазе на $\pi/2$. Амплитуда

этих колебаний равна радиусу орбиты, т. е., по порядку величины, размеру излучающего атома. Движущийся с ускорением заряд излучает электромагнитные волны, поэтому с каждым из этих колебаний связана электромагнитная волна. Так как длина волны велика по сравнению с размером атома, т. е. с амплитудой колебаний электрона, то для описания излучения атома можно воспользоваться результатами, полученными при рассмотрении излучения движущегося заряда.

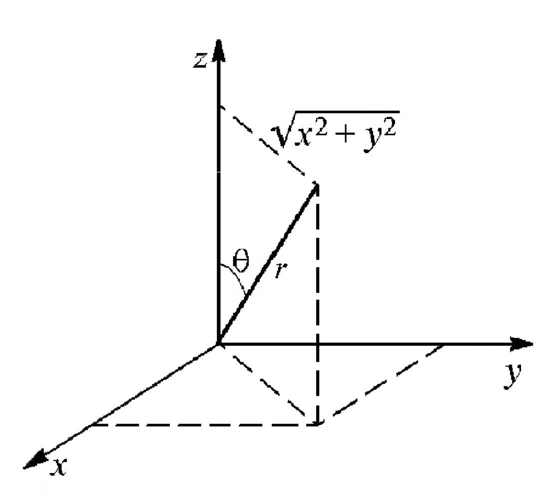


Рис. 94. К усреднению по направлениям плотности потока излучаемой энергии

Обе излучаемые вращающимся электроном волны поляризованы во взаимно перпендикулярных направлениях, они не интерферируют, и поэтому их можно рассматривать отдельно. Найдем мощность, уносимую каждой из этих волн. Среднее по времени значение плотности потока энергии волны на расстоянии r от атома дается формулой (10) § 28 кн. 2.

Так как в рассматриваемом случае величина заряда q равна заряду электрона e, то в этой формуле q^2 следует заменить на e^2, после чего она примет вид

$$\langle j \rangle = \frac{1}{8\pi} \frac{e^2 \omega^4 A^2}{c^3 r^2} \sin^2 \theta. \tag{3}$$

Чтобы найти полную мощность, излучаемую по всем направлениям, можно усреднить (3) по углу θ и умножить полученный результат на $4\pi r^2$. Как видно из рис. 94,

$$\sin^2 \theta = \frac{x^2 + y^2}{r^2}. \tag{4}$$

Так как $\langle x^2 \rangle = \langle y^2 \rangle = \langle z^2 \rangle = \frac{1}{3} r^2$, усреднение выражения (4) по всем направлениям дает

$$\langle \sin^2 \theta \rangle = \frac{2}{3}.$$

Поэтому полная излучаемая осциллятором мощность P равна

$$P = \frac{1}{3} \frac{e^2 \omega^4 A^2}{c^3}. \tag{5}$$

Излучаемая атомом мощность вдвое больше. Заменяя в (5) амплитуду колебаний A на боровский радиус a_0 и отбрасывая несущественный при оценках множитель 2/3, получаем для излучаемой атомом мощности P следующее приближенное выражение:

$$P \approx \frac{e^2 \omega^4 a_0^2}{c^3}. \tag{6}$$

Время жизни возбужденного атома. Теперь мы можем оценить промежуток времени τ, в течение которого продолжается единичный акт излучения, т. е. испускается один фотон с энергией $\hbar\omega$:

$$\tau = \frac{\hbar\omega}{P} = \frac{\hbar c^3}{e^2\omega^3 a_0^2}\,. \qquad (7)$$

С точки зрения классической физики атом начинает излучать сразу же после того, как он попадает в возбужденное состояние, и этот процесс длится в течение времени τ. По квантовой теории излучение фотона и переход атома в основное состояние происходит скачком. В какой именно момент времени это произойдет, неизвестно. Момент самопроизвольного испускания фотона есть случайная величина, а время τ можно понимать как среднее время жизни атома в возбужденном состоянии.

Из формулы (7) видно, что обусловленное спонтанным излучением время жизни атома в возбужденном состоянии обратно пропорционально кубу частоты излучения.

Рассмотрим произведение $\omega\tau$, которое с точностью до множителя 2π дает отношение времени жизни возбужденного атома к периоду колебаний. Согласно классическим представлениям $\omega\tau$ определяет число колебаний за время высвечивания. Другими словами, это есть число волн в отдельном излучаемом атомом цуге. С помощью (7) имеем

$$\omega\tau \approx \frac{\hbar c^3}{e^2\omega^2 a_0^2}\,.$$

Произведение ωa_0 равно скорости v электрона на орбите в атоме. В боровской модели атома v/c равно постоянной тонкой структуры α. Поэтому выражение для $\omega\tau$ принимает вид

$$\omega\tau \approx \frac{\hbar c}{e^2}\frac{1}{\alpha^2} = \frac{1}{\alpha^3}\,. \qquad (8)$$

Так как $\alpha = 1/137$, то $\omega\tau \approx 10^7$. Для оптических частот $\omega \approx 10^{15}$ с$^{-1}$. Отсюда получаем для времени излучения оценку $\tau \approx 10^{-8}$ с.

Таким образом, в отдельном излучаемом атомом цуге содержится около 10 миллионов волн. Пространственная протяженность такого цуга составляет несколько метров. Это означает, что соответствующая спонтанному излучению свободного атома спектральная линия имеет конечную естественную ширину $\Delta\omega$, связанную с длительностью отдельного цуга τ соотношением (5) § 34 кн. 2, где рассматривалась связь длительности цуга волн с интервалом соответствующих ему частот:

$$\tau\Delta\omega \approx 2\pi. \qquad (9)$$

Согласно формуле (8) отношение естественной ширины $\Delta\omega$ к частоте линии ω равно

$$\frac{\Delta\omega}{\omega} = \frac{\alpha^3}{2\pi} \approx 10^{-7}. \qquad (10)$$

Ширина спектральных линий. На опыте положение уровней энергии атома определяется по наблюдению спектральных линий. Так как линия излучения имеет конечную ширину, то частота излучения ω и, следовательно, энергия возбужденного состояния атома не имеют строго определенного значения. Ширина возбужденного уровня энергии ΔE равна произведению постоянной Планка на ширину спектральной линии $\Delta\omega$, излучаемой при переходе с этого уровня:

$$\Delta E = \hbar\Delta\omega. \tag{11}$$

Отношение ширины возбужденного уровня к расстоянию между уровнями, как видно из (10), равно всего одной десятимиллионной.

Таким образом, из-за спонтанных переходов возбужденные уровни энергии атома имеют хотя и малую, но конечную ширину, а даваемые квантовой механикой строго дискретные значения энергии этих уровней следует понимать как средние значения. Бесконечно узким является только наинизший уровень энергии, соответствующий основному состоянию атома. Возбужденные состояния атома, строго говоря, не являются стационарными. Они получаются строго дискретными тогда, когда рассматривается изолированный атом. Для получения истинной картины необходимо даже свободный атом рассматривать вместе с электромагнитным полем.

Ширина линий и соотношение неопределенностей. Уширение возбужденных уровней энергии атома вследствие спонтанных переходов находится в полном соответствии с соотношением неопределенностей Бора—Гейзенберга. Подставляя в (11) $\Delta\omega = 2\pi/\tau$ из (9), получаем

$$\Delta E\tau = 2\pi\hbar = h. \tag{12}$$

Неопределенность в значении энергии возбужденного уровня ΔE связана с неопределенностью в моменте спонтанного перехода в основное состояние τ соотношением (12). Это минимальная неопределенность значения энергии, допускаемая неравенством Бора—Гейзенберга, выражаемым формулой (4) § 7.

Уширение уровней энергии атома проявляется не только в испускании, но и в поглощении света: линии поглощения также имеют конечную ширину. Атом может перейти из основного состояния в возбужденное, поглощая свет в определенном интервале частот, причем ширина этого интервала обусловлена шириной возбужденного уровня энергии.

Уширение спектральных линий из-за столкновений. До сих пор мы рассматривали ширину спектральных линий изолированного атома, находящегося в покое относительно наблюдателя. Эта ширина является внутренним свойством атома. Наблюдаемые на опыте значения ширины спектральных линий, как правило, значительно больше их естественной ширины. Дополнительное уширение линий спектра разреженного газа обусловлено тепловым движением атомов и их столкновениями.

Если среднее время свободного движения атомов между столкновениями τ_1 меньше времени жизни атома в возбужденном состоянии τ, то эффективное время непрерывного излучения цуга волн сокращается до значения τ_1. Ширина линий $\Delta\omega$ в этом случае определяется формулой (9), в которой τ заменено на τ_1. Так как частота столкновений зависит от концентрации атомов, то уширение спектральных линий за счет столкновений уменьшается при понижении давления.

Доплеровское уширение спектральных линий. При низком давлении преобладающим становится уширение, обусловленное явлением Доплера. Из-за теплового движения излучающие атомы имеют различные проекции скорости v_x на направление наблюдения. Поэтому доплеровский сдвиг частоты излучения относительно частоты ω_0 излучения неподвижного атома $(\omega - \omega_0)/\omega_0 = v_x/c$ будет различен у разных атомов. В результате наблюдаемая спектральная линия, являющаяся суперпозицией узких естественных линий излучения многих атомов, уширяется. Интенсивность излучения $I(\omega)\,d\omega$ в спектральном интервале частот от ω до $\omega + d\omega$ пропорциональна числу атомов, излучающих свет в этом интервале. Чтобы найти число таких атомов, нужно в формуле для максвелловской функции распределения по скоростям заменить v_x на $c(\omega - \omega_0)/\omega_0$. В результате получим

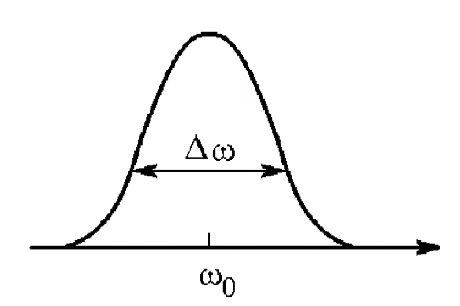


Рис. 95. Контур спектральной линии, форма которого обусловлена допплеровским уширением

$$I(\omega)\,d\omega = I_0 \exp\left[-\frac{Mc^2}{2kT}\frac{(\omega-\omega_0)^2}{\omega_0^2}\right] d\omega, \tag{13}$$

где I_0 — интенсивность излучения на частоте ω_0, а M — масса атома. Контур спектральной линии, описываемый формулой (13), показан на рис. 95. Принимая за ширину линии интервал частот $\Delta\omega$, на границах которого интенсивность в e раз меньше, чем в центре линии, получим

$$\frac{Mc^2}{2kT}\left(\frac{\Delta\omega}{2\omega_0}\right)^2 = 1,$$

откуда для доплеровской ширины линии имеем

$$\Delta\omega = \frac{2\omega_0}{c}\sqrt{\frac{2kT}{M}}\,. \tag{14}$$

Доплеровская ширина спектральных линий пропорциональна корню из термодинамической температуры газа. Так как доплеровская ширина значительно превышает естественную, то увидеть естественную ширину линии можно лишь в специальных условиях, например, наблюдая излучение пучка атомов в направлении, перпендикулярном их скорости.

- В каком соотношении находятся размер атома водорода и длина волны излучаемого им света? Каким образом в это соотношение входит постоянная тонкой структуры $\alpha = e^2/\hbar c$?
- Почему две когерентные волны, поляризованные во взаимно перпендикулярных направлениях, не интерферируют, т. е. переносимые ими энергии просто складываются?
- Как зависит излучаемая атомом мощность от частоты обращения электрона и от размера его орбиты?
- Что такое естественная ширина спектральных линий?
- Как соотношение между размером атома и длиной волны связано с малой естественной шириной спектральных линий?
- Каким образом длительность отдельного цуга волн, излучаемых атомом, связана с естественной шириной спектральной линии?
- Почему уровни энергии возбужденных состояний атома имеют конечную ширину? Чем определяется эта ширина?
- Почему в излучении света разреженными газами наблюдаемая ширина спектральных линий значительно превосходит естественную ширину?
- Поясните, каким образом максвелловское распределение атомов по скоростям проявляет себя в форме контура спектральной линии.
- Как доплеровская ширина спектральной линии зависит от температуры газа?

§ 29. Излучение света нагретыми телами

Спектральный состав излучения отдельных возбужденных атомов представляет собой набор сравнительно узких линий. Это значит, что излучаемый разреженными газами или парами свет концентрируется в узких спектральных интервалах вблизи определенных частот, характерных для атомов каждого сорта.

Тепловое излучение. Совсем иной вид имеет спектр излучения твердых и жидких тел, нагретых до высокой температуры. В этом излучении, называемом *тепловым*, присутствуют электромагнитные волны всех частот из очень широкого диапазона, т. е. его спектр является *сплошным*.

Чтобы получить представление о характере теплового излучения, рассмотрим несколько тел, нагретых до различной температуры и помещенных в замкнутую полость, внутренние стенки которой полностью отражают падающее на них излучение. Опыт показывает, что такая система, в соответствии с положениями термодинамики, рано или поздно приходит в состояние теплового равновесия, при котором все тела приобретают одинаковую температуру. Так происходит и в том случае, если внутри полости будет абсолютный вакуум и тела могут обмениваться энергией только путем излучения и поглощения электромагнитных волн. Это позволяет применить при изучении такой системы законы термодинамики.

В равновесии все тела в единицу времени поглощают столько же энергии электромагнитных волн, сколько излучают, а плотность энергии излучения, заполняющего полость, достигает некоторой определенной величины, соответствующей установившейся температуре. Такое излучение, находящееся в термодинамическом равновесии с телами, имеющими определенную температуру, называется *равновесным* или *черным* излучением. Не только плотность энергии, т. е. полная энергия единицы объема, но и спектральный состав равновесного излучения, заполняющего полость, зависит только от температуры и совершенно не зависит от свойств тел, находящихся в полости.

Спектральный состав теплового излучения. Универсальный характер спектрального состава равновесного излучения, как впервые показал Кирхгоф еще в 1860 г., непосредственно следует из второго закона термодинамики. В самом деле, предположим противное, т. е. что спектральный состав зависит от природы тела, с которым излучение находится в равновесии. Возьмем две полости, в которых излучение находится в равновесии с разными телами, имеющими, однако, одинаковую температуру. Соединим полости небольшим отверстием так, чтобы они могли обмениваться излучением. Если плотности энергии излучения в них различны, то возникает направленный перенос лучистой энергии, который приведет к самопроизвольному нарушению теплового равновесия между телами, т. е. к возникновению некоторой разности температур. Это противоречит второму закону термодинамики.

Для экспериментального изучения спектрального состава равновесного излучения можно проделать небольшое отверстие в окружающей полость оболочке. Излучение, выходящее наружу через отверстие, хотя и не является равновесным, обладает тем не менее в точности таким же спектральным составом, что и заполняющее полость равновесное излучение. Выходящее из отверстия излучение отличается от равновесного только тем, что оно не является изотропным, так как распространяется в определенном направлении.

Если увеличить температуру в полости, то будет возрастать уносимая выходящим из отверстия излучением энергия. Это означает, что объемная плотность энергии равновесного излучения растет с температурой. Этот рост происходит очень быстро, как мы увидим ниже, пропорционально четвертой степени термодинамической температуры. С увеличением температуры изменяется и спектральный состав излучения, причем таким образом, что максимум смещается в область более коротких волн: выходящий из отверстия в горячей печи свет имеет красноватый оттенок при сравнительно невысокой температуре и становится желтым и даже белым по мере ее роста.

Что можно увидеть, заглянув через отверстие внутрь полости, в которой излучение находится в равновесии с телами? Так как свойства выходящего из отверстия излучения при тепловом равновесии не зависят от природы находящихся внутри полости тел, то излучение

не может нести никакой информации об этих телах, кроме их температуры. И действительно, заглянув внутрь печи, мы не увидим ни предметов на фоне стенок полости, ни самих стенок, хотя в глаз будет попадать много света. Контуры предметов внутри полости не будут видны, все будет представляться одинаково светлым.

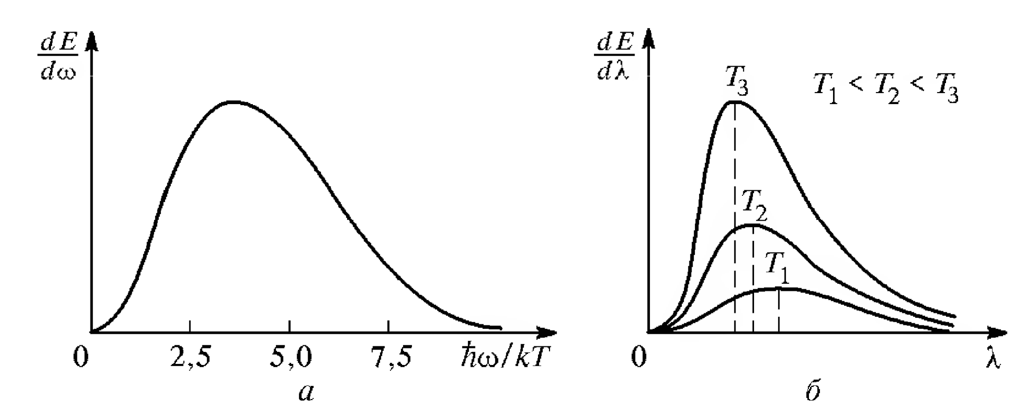


Рис. 96. Распределение энергии по частотам в спектре равновесного излучения (*а*) и спектральная плотность равновесного излучения при разных температурах (*б*)

Возможность различать предметы появляется только при использовании неравновесного излучения. Если даже это излучение исходит от раскаленных тел и его спектральный состав близок к равновесному, температура излучающей поверхности должна быть выше температуры освещаемых предметов.

Все наблюдаемые на опыте закономерности черного излучения описываются формулой Планка, полученной на основе отказа о непрерывном характере процесса излучения. Даваемое формулой Планка распределение энергии по частотам в спектре равновесного излучения

$$\frac{dE}{d\omega} = \frac{\hbar\omega^3}{\pi^2 c^3}\left(\exp\frac{\hbar\omega}{kT} - 1\right)^{-1}$$

показано на рис. 96*а*. На рис. 96*б* показана спектральная плотность равновесного излучения как функция длины волны при нескольких температурах.

Излучение как газ фотонов. Равновесное тепловое излучение можно рассматривать как газ, состоящий из фотонов. Фотонный газ является идеальным, так как разные электромагнитные волны в вакууме не взаимодействуют друг с другом. Поэтому установление теплового равновесия в фотонном газе возможно только при его взаимодействии с веществом.

Механизм установления теплового равновесия заключается в поглощении одних и испускании других фотонов веществом. Возможность поглощения и испускания фотонов приводит к характерной особенности фотонного газа: число частиц в нем не является постоянным, а само определяется из условия термодинамического равновесия.

Представление о фотонном газе позволяет очень просто найти зависимость плотности энергии равновесного излучения w от термодинамической температуры T. Это можно сделать, воспользовавшись соображениями размерности. Энергию единицы объема излучения можно представить в виде произведения среднего числа фотонов в единице объема n, равномерно заполняющих полость, на среднюю энергию одного фотона $\langle E\rangle$:

$$w = n\langle E\rangle. \tag{1}$$

Величины, от которых может зависеть средняя энергия фотона и число фотонов в единице объема равновесного излучения, — это термодинамическая температура T, постоянная Больцмана k, скорость света c и постоянная Планка $\hbar$. Поскольку равновесное излучение в полости не зависит ни от размеров и формы полости, ни от природы тел, находящихся в полости, ни от вещества ее стенок, то такие параметры, как размеры тел и полости, и такие константы, как заряды и массы электронов и ядер, не могут фигурировать в выражениях для n и $\langle E\rangle$.

Зависимость плотности энергии от температуры. Средняя энергия фотона теплового излучения по порядку величины равна kT: $\langle E\rangle = kT$. Размерность числа фотонов в единице объема n есть L^{-3}. Из величин T, k, c и $\hbar$ можно составить единственную комбинацию, имеющую размерность длины: это $c\hbar/kT$. Поэтому концентрация фотонов n пропорциональна величине $(kT/\hbar c)^3$. Подставляя это выражение в (1), можем написать

$$w = C_1 \frac{k^4}{\hbar^3 c^3} T^4, \tag{2}$$

где C_1 — некоторый безразмерный множитель.

Формула (2) показывает, что объемная плотность энергии равновесного излучения пропорциональна четвертой степени температуры в полости. Такой быстрый рост плотности энергии с температурой обусловлен не столько ростом средней энергии фотонов (которая пропорциональна T), сколько увеличением числа фотонов в полости, которое пропорционально кубу температуры.

Если в стенке полости имеется небольшое отверстие, то поток энергии излучения j через единицу площади отверстия пропорционален произведению плотности энергии в полости на скорость света c:

$$j = C_2 \frac{k^4}{\hbar^3 c^2} T^4 = \sigma T^4, \tag{3}$$

где σ — носит название постоянной Стефана—Больцмана. Точный расчет, основанный на применении статистической механики к фотонному газу, дает для нее значение, равное

$$\sigma = \frac{\pi^2 k^4}{60\hbar^3 c^2} = 5{,}67\cdot 10^{-8}\ \frac{\text{Вт}}{\text{м}^2\cdot\text{К}^4}\ .$$

Таким образом, полная интенсивность излучения из отверстия пропорциональна четвертой степени термодинамической температуры в полости.

Излучение с поверхности нагретых тел отличается от излучения из отверстия в стенке полости. Интенсивность и спектральный состав этого излучения зависят не только от температуры, но и от свойств излучающего тела. Но во многих случаях при оценках можно считать, что эти отличия невелики.

Температура поверхности Земли. В качестве примера применения закона теплового излучения (3) рассмотрим вопрос о средней температуре земной поверхности. Будем считать, что тепловой баланс Земли определяется главным образом поглощением энергии солнечного излучения и излучением энергии в пространство, а роль процессов, идущих внутри Земли, невелика. Полный поток энергии, излучаемой Солнцем, в соответствии с (3) равен $\sigma T_C^4 \cdot 4\pi R_C^2$, где T_C — температура поверхности Солнца, а R_C — его радиус. Будем считать, что вся энергия солнечного излучения, падающая на Землю, поглощается. С помощью рис. 97 легко понять, что количество поглощаемой Землей в единицу времени энергии равно

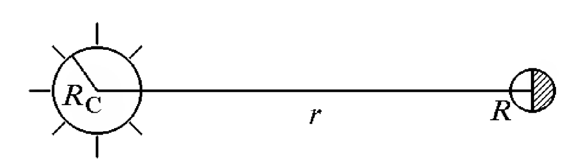


Рис. 97. К оценке средней температуры земной поверхности

$$\sigma T_C^4 \cdot 4\pi R_C^2 \frac{\pi R^2}{4\pi r^2}. \tag{4}$$

Такое же количество энергии Земля должна излучать, иначе ее температура не будет оставаться постоянной. Но излучаемая с поверхности Земли энергия также находится с помощью формулы (3), в которую в этом случае следует подставить температуру земной поверхности T. Приравнивая поток излучаемой Землей энергии $\sigma T^4 \cdot 4\pi R^2$ поглощаемой ею энергии (4), найдем

$$T = T_C \cdot \frac{1}{2} \sqrt{\frac{2R_C}{r}}. \tag{5}$$

Из этого выражения видно, что для определения T необходимо знать только температуру поверхности Солнца и угловой диаметр Солнца $2R_C/r$, видимый с Земли. Этот диаметр равен 0,01 радиана, а температура поверхности Солнца составляет примерно 6000 К. По формуле (5) находим $T \approx 300$ К.

В заключение отметим, что спектр излучения нагретых тел является настолько широким, что коэффициент полезного действия ламп накаливания и других осветительных приборов, основанных на излучении раскаленных тел, совершенно ничтожен. Область видимого света соответствует лишь узкой полосе в спектре теплового излучения.

- Почему плотность энергии и спектральный состав равновесного излучения, заполняющего полость, зависят только от температуры? Почему эти величины не могут зависеть от свойств тел, находящихся в полости, и от материала ее стенок?
- Почему выходящее наружу из отверстия в полости излучение, не являясь равновесным, имеет тем не менее тот же спектральный состав, что и равновесное излучение внутри полости? Ведь молекулы газа, вылетающие наружу через отверстие в стенке сосуда, в среднем имеют бóльшую энергию, чем молекулы в сосуде.
- Почему, заглянув через отверстие внутрь раскаленной печи, мы не увидим четких контуров находящихся там предметов?
- Почему излучение в полости, т. е. совокупность находящихся там фотонов, можно рассматривать как идеальный газ?
- Почему для установления термодинамического равновесия в газе фотонов необходимо взаимодействие фотонов с веществом?
- Как концентрация фотонов в равновесном излучении зависит от температуры?
- Как с помощью соображений размерности показать, что испускаемая телом энергия теплового излучения пропорциональна четвертой степени термодинамической температуры тела?
- Если вся приходящая на Землю от Солнца энергия в конечном счете излучается в пространство, то какой смысл имеет утверждение, что Солнце дарует жизнь всему сущему на Земле?

§ 30. Вынужденное излучение. Квантовые усилители и генераторы света

Спонтанное излучение возбужденного атома с точки зрения квантовой теории происходит скачком. Предсказать момент испускания атомом фотона невозможно, можно говорить только о среднем времени жизни τ атома в возбужденном состоянии.

Вероятность спонтанного перехода и время жизни. Пусть теперь мы имеем совокупность очень большого числа атомов, которые образуют настолько сильно разреженный газ, что взаимодействием между атомами можно пренебречь. Каждый атом может находиться в различных квантовых состояниях, которым соответствует дискретный ряд уровней энергии. Будем для простоты считать, что атом может находиться только в двух состояниях: основном с энергией E_1 и возбужденном с энергией E_2. Если какой-нибудь атом в момент времени t находится в возбужденном состоянии, то в последующий интервал времени Δt он может либо остаться в этом состоянии, либо перейти скачком в основное состояние, испустив фотон с энергией $E_2 - E_1$. Обозначим вероят-

ность спонтанного перехода атома в единицу времени из возбужденного состояния в основное через A_{21}.

Пусть в момент времени t в возбужденном состоянии находятся N_2 атомов. Так как атомы не взаимодействуют между собой, то переход в основное состояние они совершают независимо друг от друга. Поэтому число испущенных за промежуток времени Δt фотонов Δn пропорционально N_2:

$$\Delta n = N_2 A_{21} \Delta t. \tag{1}$$

Если при этом никаких процессов возбуждения атомов не происходит, то изменение числа возбужденных атомов ΔN_2, согласно (1), равно

$$\Delta N_2 = -N_2 A_{21} \Delta t. \tag{2}$$

Вероятность спонтанного перехода в единицу времени A_{21} не зависит от времени. Поэтому решение уравнения

$$\frac{dN_2}{dt} = -N_2(t) A_{21}$$

имеет вид

$$N_2(t) = N_{20} \exp(-A_{21} t), \tag{3}$$

где N_{20} — некоторая постоянная. Полагая в (3) $t = 0$, видим, что N_{20} равно числу возбужденных атомов при $t = 0$.

Таким образом, в отсутствие внешних воздействий число возбужденных атомов убывает со временем по экспоненциальному закону. Промежуток времени τ, в течение которого число возбужденных атомов в результате спонтанного излучения уменьшается в e раз, — это есть среднее время жизни атома в возбужденном состоянии. Как видно из формулы (3),

$$\tau = 1/A_{21}. \tag{4}$$

Вероятность спонтанного излучения обратно пропорциональна времени жизни τ. Как следует из формулы (7) § 28, эта вероятность пропорциональна кубу частоты излучения.

Независимый, случайный характер процессов спонтанного излучения атомов проявляется в том, что различные атомы излучают неодновременно и независимо. Поэтому фазы, направления распространения и состояния поляризации разных цугов волн не согласованы друг с другом. Это приводит к очень важному следствию: *спонтанное излучение некогерентно.*

В присутствии электромагнитного поля, кроме спонтанного излучения, будут происходить и процессы возбуждения атомов, т. е. переходы из основного состояния в возбужденное с поглощением фотонов с энергией $\hbar\omega = E_2 - E_1$. Вероятность такого перехода пропорциональна объемной плотности энергии электромагнитного поля $w(\omega)$ на

частоте перехода и некоторому коэффициенту B_{12}, характеризующему вероятность возбуждения данного атома. Полное число переходов Δn из основного состояния в возбужденное за промежуток времени Δt пропорционально числу атомов в основном состоянии N_1. Поэтому

$$\Delta n = N_1 B_{12} w(\omega)\Delta t. \quad (5)$$

Если атомы находятся в термодинамическом равновесии с электромагнитным полем, то число переходов с поглощением фотонов и с их испусканием должно быть одинаково. Основываясь на этом, можно, приравнивая (1) и (5), найти вид функции $w(\omega)$, т. е. спектральное распределение энергии равновесного излучения. Однако получающийся при этом результат не согласуется с экспериментом и с формулой Планка. Эйнштейн впервые показал, что для получения согласующегося с опытом результата необходимо предположить, что электромагнитное поле вызывает не только переходы атомов из основного состояния в возбужденное, но и переходы из возбужденного состояния в основное с излучением фотонов. Такие переходы, в отличие от спонтанных, получили название *индуцированного* или *вынужденного излучения*.

Число вынужденных переходов за время Δt пропорционально, как и число переходов с поглощением фотонов, плотности энергии электромагнитного поля на частоте перехода, числу атомов в возбужденном состоянии N_2 и некоторому коэффициенту B_{21}, характеризующему вероятность такого перехода в атоме. С учетом спонтанного излучения полное число переходов за время Δt из возбужденного состояния в основное равно

$$\Delta n = [N_2 A_{21} + N_2 B_{21} w(\omega)]\Delta t. \quad (6)$$

В состоянии термодинамического равновесия следует приравнять правые части в выражениях (5) и (6) и учесть, что отношение числа атомов в возбужденном состоянии к числу атомов в основном состоянии, в соответствии с равновесным статистическим распределением Гиббса, равно

$$\frac{N_2}{N_1} = \exp\left(-\frac{E_2 - E_1}{kT}\right) = \exp\left(-\frac{\hbar\omega}{kT}\right) \quad (7)$$

Вероятности вынужденных переходов. Легко убедиться, что вероятности B_{12} и B_{21} для атома данного сорта равны друг другу. Действительно, при очень высокой температуре, когда $kT \gg \hbar\omega$, плотность энергии w, пропорциональная четвертой степени температуры, становится настолько большой, что в формуле (6) можно пренебречь первым слагаемым по сравнению со вторым. Это означает, что в равновесии при высокой температуре вынужденное излучение преобладает над спонтанным. Приравнивая для этих условий правые части (5) и (6), имеем

$$N_1 B_{12} = N_2 B_{21} \quad (8)$$

Так как при равновесии при $kT/\hbar\omega \to \infty$, согласно (7), $N_1 \to N_2$, то из (8) получаем $B_{12} = B_{21}$.

Вероятности B_{12} и B_{21} зависят только от свойств атома и не зависят от внешних условий, в которых происходят переходы. Поэтому полученное при $T \to \infty$ равенство B_{12} и B_{21} справедливо всегда, в том числе и в отсутствие теплового равновесия.

Мазеры и лазеры. Вынужденные переходы нашли важное практическое применение. На их основе созданы квантовые генераторы излучения — мазеры, генерирующие в микроволновом диапазоне, и лазеры, излучающие в оптическом диапазоне от инфракрасных до ультрафиолетовых лучей. Наиболее существенная особенность квантовых генераторов, с которой связаны практически все их замечательные свойства, — это *когерентность* создаваемого ими излучения.

Дело в том, что волны, испущенные в результате вынужденных переходов, обладают следующей важной особенностью: их частота, фаза, направление распространения и состояние поляризации точно такие же, как и у излучения, вызвавшего переходы. Иначе говоря, фотоны, испущенные при вынужденном излучении, неотличимы от фотонов, вызывающих это излучение. Поэтому при индуцированном излучении увеличиваются только амплитуда и энергия волны.

Когерентность лазерного излучения. Пояснить когерентный характер вынужденного излучения можно следующим образом. Процесс вынужденного излучения является обратным по отношению к поглощению. При распространении пучка лучей в поглощающей среде происходит уменьшение интенсивности, но полностью сохраняются свойства когерентности. Это видно хотя бы из того, что в любых интерференционных опытах прохождение пучка света через серый фильтр, уменьшающий интенсивность без изменения спектрального состава, не разрушает интерференционной картины. Поэтому можно ожидать, что при прохождении света через среду, содержащую возбужденные атомы, в результате вынужденного излучения будет происходить усиление распространяющейся волны при сохранении ее когерентности.

Усиление света активной средой. Пусть параллельный пучок монохроматического излучения частоты ω, соответствующей разности каких-либо двух уровней энергии атомов среды, распространяется сквозь эту среду. Изменение числа фотонов в потоке на протяжении расстояния $\Delta x = c\Delta t$ за счет процессов поглощения и вынужденного излучения равно

$$\Delta n = [N_2 B_{21} w(\omega) - N_1 B_{12} w(\omega)] \frac{\Delta x}{c}. \qquad (9)$$

Вкладом спонтанного излучения в рассматриваемый поток фотонов можно пренебречь, так как спонтанное излучение распространя-

ется по всем направлениям и в направлении рассматриваемого пучка окажется ничтожная его часть. Так как $B_{12} = B_{21}$, то (9) можно переписать в виде

$$\Delta n = (N_2 - N_1) B_{12} w(\omega) \frac{\Delta x}{c}. \tag{10}$$

Эта формула позволяет понять основную идею работы квантового усилителя. Если число атомов в возбужденном состоянии N_2 меньше числа атомов в основном состоянии N_1, то при распространении волны поглощение будет преобладать над вынужденным излучением и интенсивность волны будет убывать. Так обычно и обстоит дело, если волна распространяется в среде, находившейся в состоянии термодинамического равновесия или близком к нему.

Но если в силу каких-либо причин число возбужденных атомов N_2 превосходит число атомов в основном состоянии N_1, то по мере распространения волны число фотонов в пучке будет нарастать — волна будет усиливаться.

Найдем закон, по которому будет изменяться интенсивность волны по мере ее распространения в среде. Произведение объемной плотности энергии волны $w(\omega)$ на скорость света c есть плотность потока энергии, переносимой волной:

$$j = c w(\omega).$$

Так как изменение потока фотонов на протяжении отрезка Δx равно $c\Delta n$, то изменение потока энергии Δj на этом же отрезке Δx равно $\hbar\omega c\Delta n$. Умножая обе части равенства (10) на $\hbar\omega c$, найдем

$$\Delta j = \frac{\hbar\omega}{c} (N_2 - N_1) B_{12} j \Delta x. \tag{11}$$

Переходя к пределу при $\Delta x \to 0$, получим уравнение для функции $j(x)$ — плотности потока энергии — в виде

$$\frac{dj}{dx} = \alpha j(x). \tag{12}$$

где зависящий от x коэффициент α равен

$$\alpha = \frac{\hbar\omega}{c} (N_2 - N_1) B_{12}. \tag{13}$$

Решение уравнения (12) имеет вид

$$j(x) = j_0 e^{\alpha x}. \tag{14}$$

Здесь j_0 есть плотность потока энергии пучка при $x = 0$. При $\alpha > 0$ поток энергии нарастает по мере распространения волны, а при $\alpha < 0$ убывает. Если $N_2 > N_1$, то $\alpha > 0$. В этом случае среда называется *активной*.

Методы создания активной среды. Каким же образом удается получить сильно неравновесное состояние вещества, в котором $N_2 > N_1$. В зависимости от типа рабочего вещества этого можно добиться разными методами. В импульсных твердотельных квантовых генераторах, например в рубиновом лазере, используется оптическая накачка светом мощной импульсной лампы — вспышки. В полупроводниковых лазерах непрерывного действия неравновесное состояние достигается при пропускании электрического тока через p—n-переход. В газовых лазерах атомы рабочего вещества возбуждаются электрическим разрядом.

Резонаторы лазера. В квантовых генераторах для получения монохроматического когерентного излучения, помимо эффекта усиления волны при прохождении через активную среду, используется положительная обратная связь. Часть излучаемой световой энергии все время должна оставаться внутри рабочего вещества, вызывая вынужденное излучение света все новыми возбужденными атомами. Это осуществляется с помощью зеркал. Обычно рабочее вещество помещается между двумя параллельными зеркалами, одно из которых полупрозрачно (рис. 98). Такую систему называют *оптическим резонатором*.

Возникшая в каком-либо месте в результате спонтанного перехода световая волна усиливается за счет вынужденного испускания при распространении через активную среду. Эффективно будут усиливаться только те волны, направление распространения которых совпадает с осью резонатора, так как при всех других направлениях волна быстро покинет пределы активной среды. Дойдя до полупрозрачного зеркала, волна частично выходит наружу, а частично отражается назад. Отразившаяся волна дает начало новой лавине фотонов. Пройдя вдоль резонатора от одного зеркала до другого путь L через активную среду, волна в соответствии с формулой (14) усилится в $e^{\alpha L}$ раз. Отразившаяся от второго зеркала волна снова на длине резонатора усиливается в $e^{\alpha L}$ раз и т. д. Таким образом, наличие зеркал увеличивает эффективное расстояние, которое распространяющаяся вдоль оси резонатора волна проходит в активной среде.

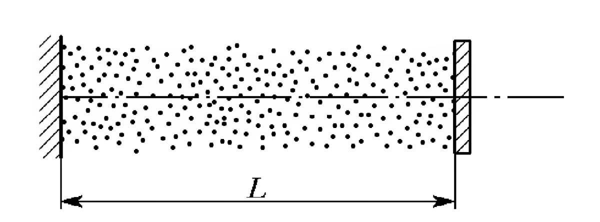


Рис. 98. Оптический резонатор с активной средой

Очевидно, что генерация возможна только тогда, когда падающая на полупрозрачное зеркало после очередного прохода волна имеет энергию не меньшую, чем при предыдущем падении. Это значит, что усиление света в активной среде должно быть достаточно большим, превышающим некоторое значение, называемое *пороговым*.

Формирование когерентного излучения. Оптический резонатор, образованный зеркалами, помимо осуществления необходимой для возникновения генерации положительной обратной связи, выполняет и другую важную функцию — формирует когерентное монохроматическое излучение. Для того чтобы выходящее через полупрозрачное зеркало излучение было когерентным, необходимо, чтобы составляющие его последовательные волновые цуги были согласованы друг с другом.

К полупрозрачному зеркалу подходят тождественные цуги волн, возникшие при вынужденных переходах в активной среде из единственного спонтанно испущенного цуга. Часть из них выходит наружу, часть отражается. Отразившиеся цуги проходят через резонатор туда и обратно и снова возвращаются к полупрозрачному зеркалу. Образовавшаяся в результате разность хода должна быть равна целому числу длин волн:

$$2L = n\lambda, \quad n = 1, 2, \ldots \tag{15}$$

Только в этом случае выходящее наружу излучение будет строго монохроматическим. Отсюда следует, что квантовый генератор может создавать монохроматическую волну не произвольной частоты, а лишь с дискретным набором частот ω_n, соответствующих допустимым значениям длины волны $\lambda_n = 2L/n$. Этот набор частот определяется формулой

$$\omega_n = \frac{2\pi c}{\lambda_n} = \frac{\pi c}{L}\, n. \tag{16}$$

Оптический резонатор, образованный зеркалами, можно рассматривать как колебательную систему, в которой собственные нормальные колебания (моды) имеют вид стоячих электромагнитных волн с узлами на зеркалах. Частоты таких колебаний определяются точно таким же условием (16). Поэтому квантовый генератор по существу представляет собой автоколебательную систему, в которой возможны незатухающие колебания на одной из собственных частот резонатора.

Спектр лазерного излучения. Возбужденный уровень энергии атомов активной среды всегда имеет конечную ширину. Поэтому усиление света при вынужденных переходах происходит не на одной фиксированной частоте, а в некотором интервале частот $\Delta\omega$, определяемом шириной спектральной линии. Обычно на интервал $\Delta\omega$ приходится несколько собственных частот резонатора (рис. 99). Некоторые из них могут оказаться в интервале $\Delta\omega_1$, в котором коэффициент усиления превосходит пороговое значение.

На этих частотах и возможна генерация излучения. Если в интервал $\Delta\omega_1$ попадает только одна собственная частота резонатора, то, казалось бы, лазер непрерывного действия должен генерировать строго монохроматический свет. Однако по ряду причин это, конечно, не так, хотя излучение такого одномодового лазера обладает

очень высокой степенью монохроматичности. И действительно, ширина линии лазерного излучения может быть много меньше не только доплеровской, но и естественной ширины спектральных линий. Ширина лазерной линии может быть доведена до значения 10^{-14} от самой частоты. Соответствующая такой монохроматичности длина цуга когерентных волн достигает сотен километров.

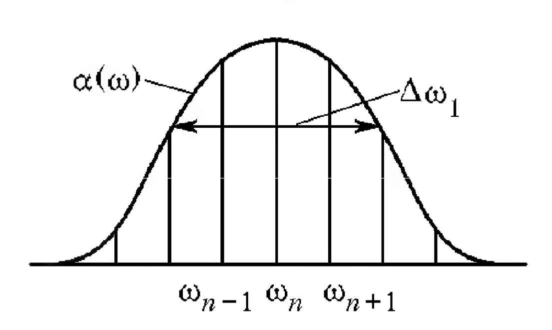


Рис. 99. В пределах контура спектральной линии находится несколько собственных частот резонатора

Наиболее существенной причиной нестабильности частоты генерации являются случайные изменения длины резонатора L, вызываемые, например, тепловым расширением, вибрациями и т. п.

Замечательной чертой лазеров, тесно связанной с когерентностью создаваемого ими излучения, является исключительная способность к *концентрации световой энергии*: к концентрации в спектре — очень узкая спектральная линия излучения, к концентрации во времени — возможность получать сверхкороткие импульсы света, к концентрации в пространстве и по направлению распространения — возможность получать практически параллельный пучок, расходимость которого определяется только дифракционными эффектами и поэтому очень мала, и сфокусировать все излучение в малой области с размерами порядка длины волны.

- Как связано среднее время жизни атома в возбужденном состоянии с классическим представлением о непрерывном излучении и с квантовым представлением о скачкообразном переходе атома из возбужденного состояния в основное?
- Как время жизни τ атома в возбужденном состоянии связано с вероятностью спонтанного перехода?
- Что такое вынужденное излучение? Какие теоретические соображения привели к выводу о его существовании?
- Как можно доказать, что вероятность поглощения фотона и вероятность вынужденного излучения одинаковы для данного перехода в атоме?
- Чем вынужденное излучение отличается от спонтанного? Почему электромагнитные волны вызывают переходы не только с поглощением энергии волны, но и с испусканием излучения?
- Чем обусловлена когерентность испускаемого света при вынужденном излучении?
- Почему при анализе работы квантового усилителя можно не учитывать спонтанное излучение?
- Сформулируйте основные этапы вывода закона нарастания интенсивности света по мере его распространения в активной среде.

- Объясните, каким образом благодаря двухзеркальному резонатору в лазере осуществляется положительная обратная связь.
- Какова роль зеркал резонатора в формировании когерентного излучения лазера?
- Чем объясняется возможность генерации излучения лазером сразу на нескольких модах резонатора?
- Объясните, почему спектральная линия лазерного излучения может иметь ширину, много меньшую естественной ширины спектральной линии атома.

§ 31. Электрический ток в газах

В обычных условиях газы не проводят электрический ток, так как их молекулы электрически нейтральны. Например, сухой воздух — это хороший изолятор, в чем мы могли убедиться с помощью самых простых опытов по электростатике. Однако воздух и другие газы становятся проводниками электрического тока, если в них тем или иным способом создать ионы. Простейший опыт, иллюстрирующий прово-

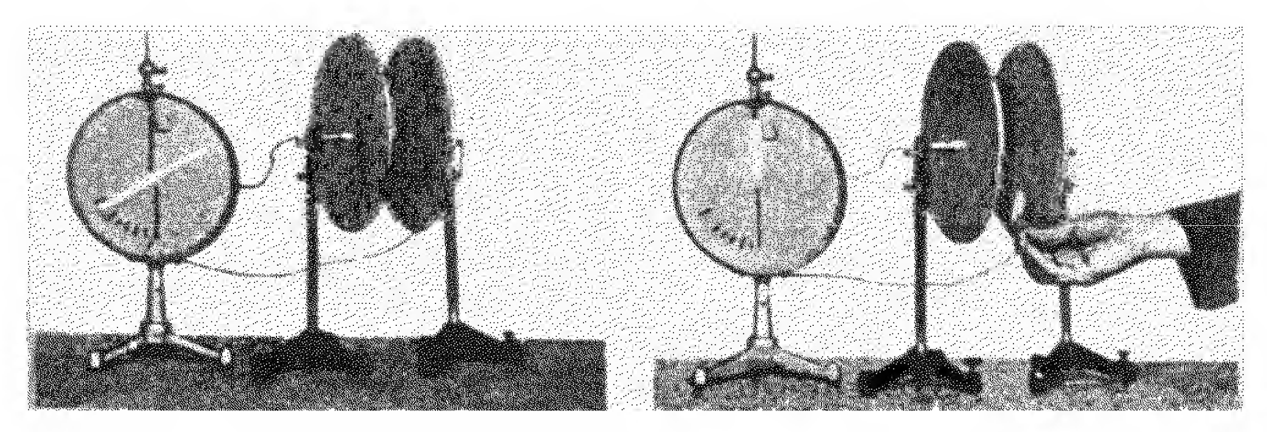


Рис. 100. Воздух становится проводником электрического тока, если его ионизировать

димость воздуха при его ионизации пламенем показан на рис. 100: заряд на пластинах, сохраняющийся в течение длительного времени, быстро исчезает при внесении зажженной спички в пространство между пластинами.

Газовый разряд. Процесс прохождения электрического тока через газ обычно называют *газовым* разрядом (или электрическим разрядом в газе). Газовые разряды подразделяются на два вида: самостоятельные и несамостоятельные.

Несамостоятельный разряд. Разряд в газе называют *несамостоятельным*, если для его поддержания необходим внешний источник ионизации. Ионы в газе могут возникать под действием высоких температур, рентгеновского и ультрафиолетового излучения, радиоактивности, космических лучей и т. д. Во всех этих случаях происходит освобождение одного или нескольких электронов из электрон-

ной оболочки атома или молекулы. В результате в газе появляются положительные ионы и свободные электроны. Освободившиеся электроны могут присоединяться к нейтральным атомам или молекулам, превращая их в отрицательные ионы.

Ионизация и рекомбинация. Наряду с процессами *ионизации* в газе происходят и обратные процессы *рекомбинации*: соединяясь между собой, положительные и отрицательные ионы или положительные ионы и электроны образуют нейтральные молекулы или атомы.

Изменение со временем концентрации n ионов, обусловленное постоянным источником ионизации и процессами рекомбинации, можно описать следующим образом. Допустим, что источник ионизации создает в единице объема газа за единицу времени q положительных ионов и такое же число электронов. Если в газе нет электрического тока и можно пренебречь уходом ионов из рассматриваемого объема из-за диффузии, то единственным механизмом уменьшения концентрации ионов n будет рекомбинация.

Рекомбинация происходит при встрече положительного иона с электроном. Число таких встреч пропорционально как числу ионов, так и числу свободных электронов, т. е. пропорционально n^2. Поэтому убыль числа ионов в единице объема в единицу времени может быть записана в виде αn^2, где α постоянная величина, называемая коэффициентом рекомбинации.

При справедливости введенных предположений уравнение баланса ионов в газе запишется в виде

$$\frac{dn}{dt} = q - \alpha n^2. \tag{1}$$

Мы не будем решать это дифференциальное уравнение в общем виде, а рассмотрим некоторые интересные частные случаи.

Прежде всего отметим, что процессы ионизации и рекомбинации через некоторое время должны скомпенсировать друг друга и в газе установится постоянная концентрация n_∞. Из (1) видно, что при $dn/dt = 0$

$$n_\infty = \sqrt{\frac{q}{\alpha}} \tag{2}$$

Стационарная концентрация ионов n_∞ тем больше, чем мощнее источник ионизации и чем меньше коэффициент рекомбинации α.

После выключения ионизатора убывание концентрации ионов $n(t)$ описывается уравнением (1), в котором нужно положить $q = 0$, а n_∞ принять в качестве начального значения концентрации $n(0)$:

$$\frac{dn}{dt} = -\alpha n^2.$$

Переписав это уравнение в виде $\frac{dn}{n^2} = -\alpha dt$, после интегрирования получаем

$$\frac{1}{n(t)} - \frac{1}{n_\infty} = \alpha t,$$

откуда

$$n(t) = \frac{n_\infty}{1 + n_\infty \alpha t}. \tag{3}$$

График этой функции показан на рис. 101. Он представляет собой гиперболу, асимптотами которой являются ось времени и вертикальная прямая $t = 1/(\alpha n_\infty)$. Разумеется, физический смысл имеет лишь участок гиперболы, соответствующий значениям $t > 0$. Отметим медленный характер убывания концентрации со временем в сравнении с часто встречающимися в физике процессами экспоненциального затухания, которые реализуются, когда скорость убывания какой-либо величины пропорциональна первой степени мгновенного значения этой величины.

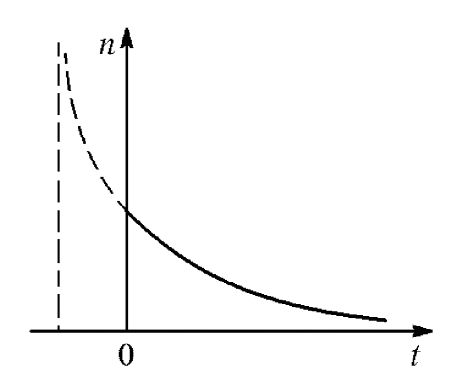


Рис. 101. Убывание концентрации ионов в газе после выключения источника ионизации

Несамостоятельная проводимость. Процесс спадания концентрации ионов после прекращения действия ионизатора значительно ускоряется, если газ находится во внешнем электрическом поле. Вытягивая электроны и ионы на электроды, электрическое поле может очень быстро обратить в нуль электропроводность газа в отсутствие ионизатора.

Для уяснения закономерностей несамостоятельного разряда рассмотрим для простоты случай, когда ток в ионизуемом внешним источником газе течет между двумя плоскими электродами, параллельными друг другу. В этом случае ионы и электроны находятся в однородном электрическом поле напряженности E, равной отношению приложенного к электродам напряжения U к расстоянию l между ними.

Подвижность электронов и ионов. При постоянном приложенном напряжении в цепи устанавливается некоторая постоянная сила тока I. Это значит, что электроны и ионы в ионизованном газе движутся с постоянными скоростями. Чтобы объяснить этот факт, нужно считать, что кроме постоянной ускоряющей силы электрического поля на движущиеся ионы и электроны действуют силы сопротивления, растущие с увеличением скорости. Эти силы описывают усредненный эффект столкновений электронов и ионов с нейтральными атомами и молекулами газа. Благодаря силам сопротивления устанавливаются в среднем постоянные скорости $\mathbf{v}_e$ и $\mathbf{v}_i$ электронов и ионов, пропорциональные напряженности E электрического поля:

$$\mathbf{v}_e = -b_e \mathbf{E}, \qquad \mathbf{v}_i = b_i \mathbf{E}. \tag{4}$$

Коэффициенты пропорциональности b_e и b_i называются *подвижностями* электрона и иона. Подвижности ионов и электронов имеют разные значения и зависят от сорта газа, его плотности, температуры и т. д.

Плотность электрического тока **j**, т. е. заряд, переносимый электронами и ионами за единицу времени через единичную площадку, выражается через концентрацию электронов n_e и ионов n_i, их заряды $-e$ и e_i и скорости установившегося движения $\mathbf{v}_e$ и $\mathbf{v}_i$:

$$\mathbf{j} = -en_e\mathbf{v}_e + e_in_i\mathbf{v}_i. \qquad (5)$$

Квазинейтральность. В обычных условиях ионизованный газ в целом электронейтрален, или, как говорят, *квазинейтрален*, ибо в малых объемах, содержащих сравнительно небольшое число электронов и ионов, условие электронейтральности может и нарушаться. Это значит, что выполняется соотношение

$$-en_e + e_in_i = 0. \qquad (6)$$

Будем для простоты считать, что ионы однозарядные. В этом случае концентрации ионов и электронов одинаковы: $n_e = n_i = n$. Тогда с помощью (4) выражснис для плотности тока (5) можно записать в виде

$$\mathbf{j}=ne(-\mathbf{v}_e + \mathbf{v}_i) = ne(b_e + b_i)\mathbf{E}. \qquad (7)$$

Плотность тока при несамостоятельном разряде. Чтобы получить закон изменения со временем концентрации носителей тока при несамостоятельном разряде в газе, нужно наряду с процессами ионизации внешним источником и рекомбинации учесть также уход электронов и ионов на электроды. Число частиц, уходящих в единицу времени на электрод площади S из объема Sl, равно jS/e. Скорость $\frac{dn}{dt}$ убывания концентрации $n(t)$ таких частиц мы получим, разделив это число на объем Sl газа между электродами. Поэтому уравнение баланса вместо (1) при наличии тока запишется в виде

$$\frac{dn}{dt} = q - \alpha n^2 - \frac{j}{el}. \qquad (8)$$

Для установления режима, когда $\frac{dn}{dt} = 0$, из (8) получаем

$$q = \alpha n^2 + \frac{j}{el}. \qquad (9)$$

Уравнение (9) позволяет найти зависимость плотности j установившегося тока при несамостоятельном разряде от приложенного напряжения U (или от напряженности поля E).

Два предельных случая видны непосредственно.

Закон Ома. При низком напряжении, когда $j \to 0$, в уравнении (9) можно пренебречь вторым слагаемым в правой части, после чего получаем $n = \sqrt{\frac{q}{\alpha}}$. Из формулы (7) при этом имеем

$$\mathbf{j} = e\sqrt{\frac{q}{\alpha}}\,(b_e + b_i)\mathbf{E}. \tag{10}$$

Плотность тока **j** пропорциональна напряженности приложенного электрического поля. Таким образом, для несамостоятельного газового разряда в слабых электрических полях выполняется закон Ома.

Ток насыщения. При низкой концентрации n электронов и ионов в уравнении (9) можно пренебречь первым (квадратичным по n) слагаемым в правой части. В этом приближении вектор плотности тока направлен вдоль напряженности электрического поля, а его модуль

$$j = qel \tag{11}$$

не зависит от приложенного напряжения. Этот результат справедлив для сильных электрических полей. В этом случае говорят о токе насыщения.

Оба рассмотренных предельных случая можно исследовать и не обращаясь к уравнению (9). Однако таким путем нельзя проследить, как при увеличении напряжения происходит переход от закона Ома к нелинейной зависимости тока от напряжения.

В первом предельном случае, когда ток очень мал, основной механизм удаления электронов и ионов из области разряда — это рекомбинация. Поэтому для стационарной концентрации n можно воспользоваться выражением (2), что при учете (7) немедленно дает формулу (10). Во втором предельном случае, наоборот, пренебрегается рекомбинацией. В сильном электрическом поле электроны и ионы не успевают сколько-нибудь заметно рекомбинировать за время пролета от одного электрода до другого, если концентрация их достаточно мала. Тогда все образуемые внешним источником электроны и ионы достигают электродов и полная плотность тока равна qel. Она пропорциональна длине l ионизационной камеры, поскольку полное число производимых ионизатором электронов и ионов пропорционально l.

Экспериментальное изучение газового разряда. Выводы теории несамостоятельного газового разряда подтверждаются экспериментами. Для исследования разряда в газе удобно использовать стеклянную трубку с двумя металлическими электродами. Электрическая схема такой установки показана на рис. 102. Подвижности электронов и ионов сильно зависят от давления газа (обратно пропорционально давлению), поэтому опыты удобно проводить при пониженном давлении.

На рис. 103 представлена зависимость силы тока I в трубке от приложенного к электродам трубки напряжения U. Ионизацию

в трубке можно создать, например, рентгеновскими или ультрафиолетовыми лучами либо с помощью слабого радиоактивного препарата. Существенно только, чтобы внешний источник ионов оставался неизменным ($q = \text{const}$). Линейный участок OA вольт-амперной характеристики соответствует области применимости закона

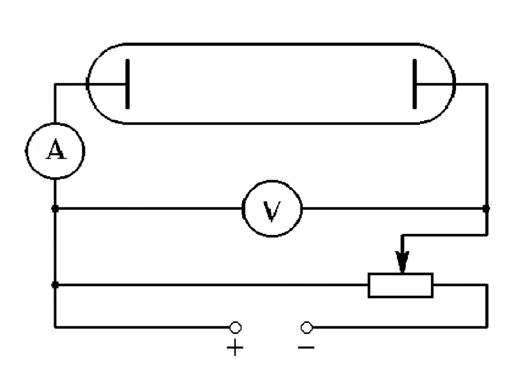


Рис. 102. Схема установки для изучения газового разряда

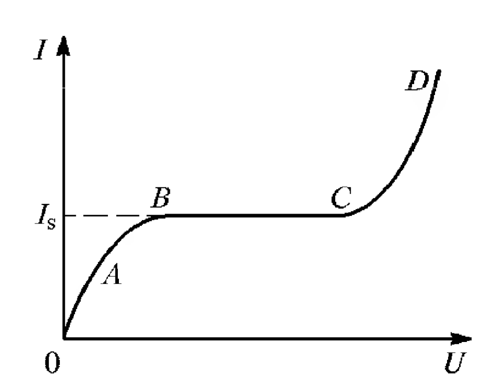


Рис. 103. Экспериментальная вольт-амперная характеристика газового разряда

Ома. На участке AB сила тока нелинейно зависит от напряжения. Начиная с точки B ток достигает насыщения и остается постоянным на некотором участке BC. Все это соответствует теоретическим предсказаниям.

Самостоятельный разряд. Однако в точке C снова начинается возрастание тока, сначала медленное, а затем очень резкое. Это означает, что в газе появился новый, внутренний источник ионов. Если теперь убрать внешний источник, то разряд в газе не прекращается, т. е. из несамостоятельного разряд переходит в *самостоятельный*. При самостоятельном разряде образование новых электронов и ионов происходит в результате внутренних процессов, происходящих в самом газе.

Ионизация электронным ударом. Нарастание тока при переходе от несамостоятельного разряда к самостоятельному происходит лавинообразно и называется *электрическим пробоем газа*. Напряжение, при котором происходит пробой, называется напряжением зажигания. Оно зависит от рода газа и от произведения давления газа p на расстояние l между электродами.

Процессы в газе, ответственные за лавинообразное нарастание силы тока при увеличении приложенного напряжения, связаны с ионизацией нейтральных атомов или молекул газа свободными электронами, разогнанными электрическим полем до достаточно больших энергий. Кинетическая энергия электрона перед очередным столкновением с нейтральным атомом или молекулой пропорциональна напряженности электрического поля E и длине свободного пробега электрона λ:

$$\frac{mv^2}{2} = eE\lambda. \tag{12}$$

Если эта энергия достаточна для того, чтобы ионизовать нейтральный атом или молекулу, т. е. превосходит работу ионизации A_i

$$eE\lambda > A_i,$$

то при столкновении электрона с атомом или молекулой происходит их ионизация. В результате вместо одного электрона возникают два. Они в свою очередь разгоняются электрическим полем и ионизуют встречающиеся на их пути атомы или молекулы и т. д. Процесс развивается лавинообразно и называется электронной лавиной. Описанный механизм ионизации называется ионизацией электронным ударом.

Экспериментальное доказательство того, что ионизация нейтральных атомов газа происходит в основном благодаря ударам электронов, а не положительных ионов, было дано Дж. Таунсендом. Он брал ионизационную камеру в виде цилиндрического конденсатора, внутренним электродом которого служила тонкая металлическая нить, натянутая по оси цилиндра. В такой камере ускоряющее электрическое поле сильно неоднородно, и основную роль в ионизации играют частицы, которые попадают в область наиболее сильного поля вблизи нити. Опыт показывает, что при одном и том же напряжении между электродами ток разряда больше в том случае, когда положительный потенциал подается на нить, а не на внешний цилиндр. Именно в этом случае все создающие ток свободные электроны обязательно проходят через область наиболее сильного поля.

Эмиссия электронов из катода. Самостоятельный разряд может быть стационарным лишь при условии постоянного появления в газе новых свободных электронов, так как все возникающие в лавине электроны достигают анода и выбывают из игры. Новые электроны выбиваются из катода положительными ионами, которые при движении к катоду также ускоряются электрическим полем и приобретают достаточную для этого энергию.

Катод может испускать электроны не только в результате бомбардировки ионами, но и самостоятельно, при нагревании его до высокой температуры. Такой процесс называется *термоэлектронной эмиссией*, его можно рассматривать как своего рода испарение электронов из металла. Обычно оно происходит при таких температурах, когда испарение самого материала катода еще мало. В случае самостоятельного газового разряда катод обычно разогревается не нитью накала, как в электронных лампах, а из-за выделения теплоты при бомбардировке его положительными ионами. Поэтому катод испускает электроны даже тогда, когда энергия ионов недостаточна для выбивания электронов.

Самостоятельный разряд в газе возникает не только в результате перехода от несамостоятельного при повышении напряжения и удалении внешнего источника ионизации, но и при непосредственном при-

ложении напряжения, превышающего пороговое напряжение зажигания. Теория показывает, что для зажигания разряда достаточно самого незначительного количества ионов, которые всегда присутствуют в нейтральном газе хотя бы из-за естественного радиоактивного фона.

В зависимости от свойств и давления газа, конфигурации электродов и приложенного к электродам напряжения возможны различные виды самостоятельного разряда.

Тлеющий разряд. При низких давлениях (десятые и сотые доли миллиметра ртутного столба) в трубке наблюдается *тлеющий разряд.* Для зажигания тлеющего разряда достаточно напряжения в несколько сотен или даже десятков вольт. В тлеющем разряде можно выделить четыре характерные области. Это — темное катодное пространство, тлеющее (или отрицательное) свечение, фарадеево темное пространство и светящийся положительный столб, занимающий бóльшую часть пространства между анодом и катодом.

Первые три области находятся вблизи катода. Именно здесь происходит резкое падение потенциала, связанное с большой концентрацией положительных ионов на границе катодного темного пространства и тлеющего свечения. Электроны, ускоренные в области катодного темного пространства, производят в области тлеющего свечения интенсивную ударную ионизацию. Тлеющее свечение обусловлено рекомбинацией ионов и электронов в нейтральные атомы или молекулы. Для положительного столба разряда характерно незначительное падение потенциала и свечение, вызываемое возвращением возбужденных атомов или молекул газа в основное состояние.

Коронный разряд. При сравнительно высоких давлениях в газе (порядка атмосферного) вблизи заостренных участков проводника, где электрическое поле сильно неоднородно, наблюдается разряд, светящаяся область которого напоминает корону. *Коронный разряд* иногда возникает в естественных условиях на верхушках деревьев, корабельных мачтах и т. п. («огни святого Эльма»). С коронным разрядом приходится считаться в технике высоких напряжений, когда этот разряд возникает вокруг проводов высоковольтных линий электропередачи и приводит к потерям электроэнергии. Полезное практическое применение коронный разряд находит в электрофильтрах для очистки промышленных газов от примесей твердых и жидких частиц.

При увеличении напряжения между электродами коронный разряд переходит в *искровой* с полным пробоем промежутка между электродами. Он имеет вид пучка ярких зигзагообразных разветвляющихся каналов, мгновенно пронизывающих разрядный промежуток и прихотливо сменяющих друг друга. Искровой разряд сопровождается выделением большого количества теплоты, ярким голубовато-белым свечением и сильным потрескиванием. Его можно наблюдать между шариками электрофорной машины. Пример ги-

гантского искрового разряда — естественная молния, где сила тока достигает $5 \cdot 10^5$ А, а разность потенциалов — 10^9 В.

Поскольку искровой разряд происходит при атмосферном (и более высоком) давлении, то напряжение зажигания весьма велико: в сухом воздухе при расстоянии между электродами 1 см оно составляет около 30 кВ.

Электрическая дуга. Специфическим практически важным видом самостоятельного газового разряда является *электрическая дуга*. При соприкосновении двух угольных или металлических электродов в месте их контакта выделяется большое количество теплоты из-за большого сопротивления контакта. В результате начинается термоэлектронная эмиссия и при раздвижении электродов между ними возникает ярко светящаяся дуга из сильно ионизованного хорошо проводящего газа. Сила тока даже в небольшой дуге достигает нескольких ампер, а в большой дуге — нескольких сотен ампер при напряжении порядка 50 В. Электрическая дуга широко применяется в технике как мощный источник света, в электропечах и для электросварки.

Опыты Франка и Герца. При экспериментальном изучении газового разряда Д. Франком и Г. Герцем в 1913 г. было получено доказательство существования стационарных состояний атома с дискретными значениями энергии, сыгравшее важную роль для подтверждения развивавшейся Н. Бором теории атома.

Схема установки Франка и Герца показана на рис. 104. Стеклянный сосуд с тремя электродами заполнен парами ртути при низком давлении. Батарея $Б_1$ создает ускоряющее электрическое поле, при-

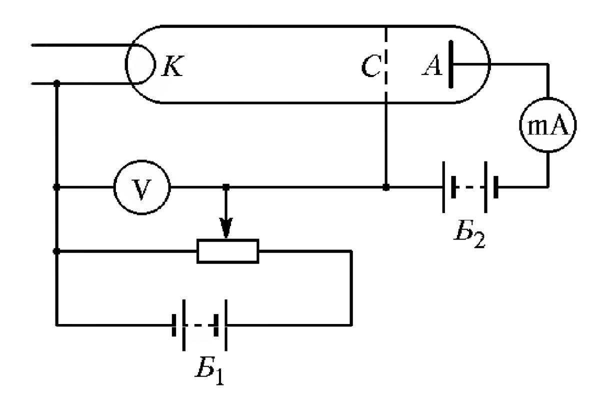


Рис. 104. Схема опытов Франка и Герца

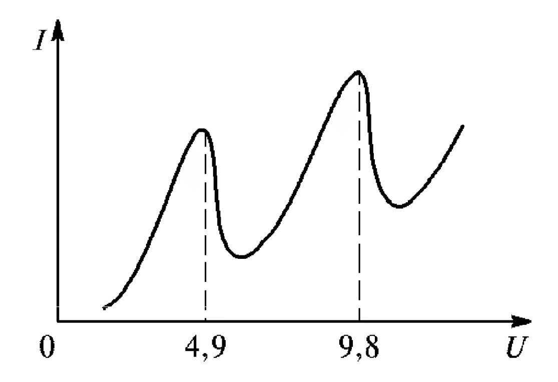


Рис. 105. Зависимость тока от ускоряющего напряжения в опытах Франка и Герца

чем напряжение U между катодом K и сеткой C можно регулировать с помощью потенциометра. Между сеткой и анодом создается с помощью батареи $Б_2$ слабое задерживающее поле с напряжением около 0,5 В. Это поле препятствует попаданию на анод медленных электронов. Электроны испускаются катодом K, подогреваемым электрическим током.

На рис. 105 показана полученная в этих опытах зависимость силы тока I в анодной цепи от ускоряющего напряжения U. Эта зависимость имеет немонотонный характер с максимумами при напряжениях U, кратных 4,9 В.

Дискретность уровней энергии атома. Объяснить такую зависимость тока от напряжения можно лишь наличием у атомов ртути дискретных стационарных состояний. Если бы дискретных стационарных состояний у атома не было, т. е. его внутренняя энергия могла бы принимать любые значения, то неупругие столкновения, сопровождающиеся увеличением внутренней энергии атома, могли бы происходить при любых энергиях электронов. Если же дискретные состояния есть, то столкновения электронов с атомами могут быть только упругими, пока энергия электронов недостаточна для перевода атома из основного состояния в наинизшее возбужденное.

При упругих столкновениях кинетическая энергия электронов практически не меняется, так как масса электрона много меньше массы атома ртути. В этих условиях число электронов, достигающих анода, монотонно увеличивается с ростом напряжения. Когда ускоряющее напряжение достигает значения 4,9 В, столкновения электронов с атомами становятся неупругими. Внутренняя энергия атомов скачком увеличивается, а электрон в результате соударения теряет почти всю свою кинетическую энергию.

Задерживающее поле не пропускает также медленные электроны к аноду и сила тока резко уменьшается. Она не обращается в нуль лишь потому, что часть электронов достигает сетки, не испытав неупругих соударений. Второй и последующие максимумы силы тока получаются потому, что при напряжениях, кратных 4,9 В, электроны на пути к сетке могут испытать несколько неупругих столкновений с атомами ртути.

Итак, необходимую для неупругого соударения энергию электрон приобретает только после прохождения разности потенциалов 4,9 В. Это означает, что внутренняя энергия атомов ртути не может измениться на величину, меньшую $\Delta E = 4{,}9$ эВ, что и доказывает дискретность энергетического спектра атома. Справедливость этого вывода подтверждается еще и тем, что при напряжении 4,9 В разряд начинает светиться: возбужденные атомы при спонтанных переходах в основное состояние излучают видимый свет, частота которого совпадает с вычисленной по формуле: $\hbar\omega = \Delta E$.

В классических опытах Франка и Герца методом электронного удара были определены не только потенциалы возбуждения, но и ионизационные потенциалы ряда атомов.

- Приведите пример опыта по электростатике, из которого можно сделать вывод о том, что сухой воздух — это хороший изолятор.
- Где в технике используются изолирующие свойства воздуха?

- Что такое несамостоятельный газовый разряд? При каких условиях он протекает?
- Поясните, почему скорость убывания концентрации, обусловленная рекомбинацией, пропорциональна квадрату концентрации электронов и ионов. Почему эти концентрации можно считать одинаковыми?
- Почему для закона убывания концентрации, выражаемого формулой (3), не имеет смысла вводить понятие характерного времени, широко используемого для экспоненциально затухающих процессов, хотя и в том и в другом случае процессы продолжаются, вообще говоря, бесконечно долго?
- Как по-вашему, почему в определениях подвижностей в формулах (4) для электронов и ионов выбраны противоположные знаки?
- Как сила тока при несамостоятельном газовом разряде зависит от приложенного напряжения? Почему с ростом напряжения происходит переход от закона Ома к току насыщения?
- Электрический ток в газе осуществляется как электронами, так и ионами. Однако на каждый из электродов приходят заряды лишь одного знака. Как это согласуется с тем, что во всех участках последовательной цепи сила тока одинакова?
- Почему в ионизации газа в разряде из-за соударений наибольшую роль играют электроны, а не положительные ионы?
- Опишите характерные признаки различных видов самостоятельного газового разряда.
- Почему результаты опытов Франка и Герца свидетельствуют о дискретности уровней энергии атомов?
- Опишите физические процессы, происходящие в газоразрядной трубке в опытах Франка и Герца, при повышении ускоряющего напряжения.

§ 32. Электрический ток в жидкостях

В отношении своих электрических свойств жидкости отличаются большим многообразием. Расплавленные металлы, как и металлы в твердом состоянии, имеют высокую электропроводность, связанную с большой концентрацией свободных электронов.

Многие жидкости, например чистая вода, спирт, керосин, являются хорошими диэлектриками, поскольку их молекулы электронейтральны и в них отсутствуют свободные носители заряда.

Электролиты. Особый класс жидкостей составляют так называемые *электролиты*, к которым относятся водные растворы неорганических кислот, солей и оснований, расплавы ионных кристаллов и т. д. Для электролитов характерно наличие высоких концентраций ионов, обусловливающих возможность прохождения электрического тока. Эти ионы возникают при плавлении и при растворении, когда

под влиянием электрических полей молекул растворителя происходит разложение молекул растворяемого вещества на отдельные положительно и отрицательно заряженные ионы. Такой процесс называется *электролитической диссоциацией*.

Электролитическая диссоциация. Степень диссоциации α данного вещества, т. е. доля молекул растворенного вещества, распавшихся на ионы, зависит от температуры, концентрации раствора и диэлектрической проницаемости растворителя. С увеличением температуры степень диссоциации растет. Ионы противоположных знаков могут рекомбинировать, объединяясь снова в нейтральные молекулы. При неизменных внешних условиях в растворе устанавливается динамическое равновесие, при котором процессы рекомбинации и диссоциации компенсируют друг друга.

Качественно зависимость степени диссоциации α от концентрации n растворенного вещества можно установить с помощью следующих простых рассуждений. Если в единице объема содержится n молекул растворенного вещества, то αn из них диссоциированы, а остальные $(1-\alpha)n$ не диссоциированы. Число элементарных актов диссоциации в единице объема раствора пропорционально числу нерасщепленных молекул и поэтому равно $A(1-\alpha)n$, где A — коэффициент, зависящий от природы электролита и температуры. Число актов рекомбинации пропорционально числу соударений разноименных ионов, т. е. пропорционально числу как тех, так и других ионов. Поэтому оно равно $B\alpha^2n^2$, где B — коэффициент, постоянный для данного вещества при определенной температуре.

В состоянии динамического равновесия

$$A(1-\alpha)n = B\alpha^2n^2,$$

откуда

$$\frac{\alpha^2}{1-\alpha} = \frac{A}{B}\frac{1}{n}. \qquad (1)$$

Отношение A/B не зависит от концентрации n. Видно, что чем меньше концентрация раствора, тем α ближе к единице: в очень разбавленных растворах практически все молекулы растворенного вещества диссоциированы.

Чем выше диэлектрическая проницаемость растворителя, тем больше ослабляются ионные связи в молекулах растворенного вещества и, следовательно, тем больше степень диссоциации. Так, соляная кислота HCl дает электролит с высокой электропроводностью при растворении в воде ($\varepsilon = 81$), в то время как ее раствор в этиловом эфире ($\varepsilon = 4{,}3$) очень плохо проводит электрический ток.

Необычные электролиты. Встречаются и весьма необычные электролиты. Например, электролитом является стекло, представляющее собой сильно переохлажденную жидкость, обладающую громадной вязкостью. При нагревании стекло размягчается и его вязкость

сильно уменьшается. Присутствующие в стекле ионы натрия Na^+ приобретают заметную подвижность, и становится возможным прохождение электрического тока, хотя при обычных температурах стекло является хорошим изолятором.

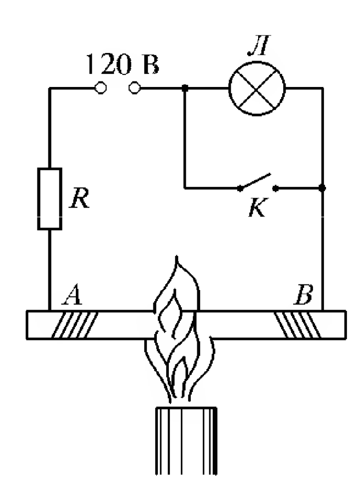


Рис. 106. Демонстрация электропроводности стекла при нагревании

Наглядной демонстрацией этого может служить опыт, схема которого показана на рис. 106. Стеклянная палочка включена в осветительную сеть через реостат *R*. Пока палочка холодная, ток в цепи ничтожный из-за высокого сопротивления стекла. Если палочку нагреть газовой горелкой до температуры 300–400 °C, то ее сопротивление упадет до нескольких десятков омов, и нить лампочки раскалится. Теперь можно закоротить лампочку ключом *K*. При этом сопротивление цепи уменьшится и сила тока возрастет. В таких условиях палочка будет эффективно нагреваться электрическим током и раскаляться до яркого свечения, даже если убрать горелку.

Ионная проводимость. Прохождение электрического тока в электролите описывается законом Ома

$$\mathbf{j} = \sigma \mathbf{E}. \qquad (2)$$

Электрический ток в электролите возникает при сколь угодно малом приложенном напряжении.

Носителями заряда в электролите являются положительно и отрицательно заряженные ионы. Механизм электропроводности электролитов во многом сходен с описанным выше механизмом электропроводности газов. Основные отличия связаны с тем, что в газах сопротивление движению носителей заряда обусловлено главным образом их столкновениями с нейтральными атомами. В электролитах подвижность ионов обусловлена внутренним трением — вязкостью — при их движении в растворителе.

При повышении температуры проводимость электролитов, в противоположность металлам, увеличивается. Это связано с тем, что с ростом температуры возрастает степень диссоциации и уменьшается вязкость.

В отличие от электронной проводимости, характерной для металлов и полупроводников, где прохождение электрического тока не сопровождается каким бы то ни было изменением химического состава вещества, ионная проводимость связана с переносом вещества и выделением на электродах веществ, входящих в состав электролитов. Такой процесс называется *электролизом*.

Электролиз. При выделении вещества на электроде уменьшается концентрация соответствующих ионов в примыкающей к электроду области электролита. Тем самым здесь нарушается динамическое равновесие между диссоциацией и рекомбинацией: именно здесь происходит разложение вещества в результате электролиза.

Электролиз впервые наблюдался в 1800 г. при разложении воды током от вольтова столба. Через несколько лет знаменитый химик Г. Дэви открыл натрий, выделив его путем электролиза из едкого натра. Количественные законы электролиза были экспериментально установлены М. Фарадеем в 1833 г. Их легко обосновать, исходя из механизма явления электролиза.

Законы Фарадея. Каждый ион обладает электрическим зарядом, кратным элементарному заряду e. Другими словами, заряд иона равен νe, где ν — целое число, равное валентности соответствующего химического элемента или соединения. Пусть при прохождении тока на электроде выделилось N ионов. Их заряд по абсолютной величине равен $N\nu e$. Положительные ионы достигают катода и их заряд нейтрализуется электронами, притекающими к катоду по проводам от источника тока. Отрицательные ионы подходят к аноду и такое же количество электронов уходит по проводам к источнику тока. При этом по замкнутой электрической цепи проходит заряд $q = N\nu e$.

Обозначим через m массу вещества, выделившегося на одном из электродов, а через m_0 массу иона (атома или молекулы). Очевидно, что $m = m_0 N$ и, следовательно, $m = m_0 q/(\nu e)$. Умножив числитель и знаменатель этой дроби на постоянную Авогадро N_A, получим

$$m = \frac{M}{\nu}\frac{q}{F}. \tag{3}$$

где $M = N_A m_0$ — атомная или молярная масса, а F — *постоянная Фарадея*, определяемая выражением

$$F = N_A e. \tag{4}$$

Из (4) видно, что постоянная Фарадея имеет смысл «одного моля электричества», т. е. это суммарный электрический заряд одного моля элементарных зарядов: $F = 96485$ Кл/моль.

Формула (3) содержит оба закона Фарадея. Она говорит о том, что масса выделившегося при электролизе вещества пропорциональна прошедшему по цепи заряду (первый закон Фарадея):

$$m = kq. \tag{5}$$

Коэффициент $k = M/(\nu F) = m_0/(\nu e)$ называется электрохимическим эквивалентом данного вещества и выражается в килограммах на кулон (кг/Кл). Он имеет смысл обратной величины удельного заряда иона. Электрохимический эквивалент k пропорционален химическому эквиваленту вещества M/ν (второй закон Фарадея).

Законы Фарадея и элементарный заряд. Поскольку во времена Фарадея представления об атомарной природе электричества еще не существовало, экспериментальное открытие законов электролиза было далеко не тривиальным. Напротив, именно законы Фарадея послужили по существу первым экспериментальным доказательством справедливости этих представлений.

Измерение на опыте постоянной Фарадея позволило в 1874 г. впервые получить числовую оценку значения элементарного заряда задолго до прямых измерений элементарного электрического заряда в опытах Милликена с масляными каплями. Замечательно, что идея атомарной структуры электричества получила недвусмысленное экспериментальное подтверждение в опытах по электролизу, выполненных в 30-е годы XIX века, когда даже идея атомарного строения вещества еще не разделялась всеми учеными. В знаменитой речи, произнесенной в Королевском обществе и посвященной памяти Фарадея, Гельмгольц таким образом комментировал это обстоятельство:

«Если мы признаем существование атомов химических элементов, то мы не можем избежать и дальнейшего заключения, что электричество, как положительное, так и отрицательное, разделено на определенные элементарные количества, которые ведут себя как атомы электричества».

Химические источники тока. Если какой-либо металл, например цинк, погрузить в воду, то некоторое количество положительных ионов цинка Zn^{++} под влиянием полярных молекул воды начнет переходить из поверхностного слоя кристаллической решетки металла в воду. В результате цинк зарядится отрицательно, а вода положительно. На границе металла и воды образуется тонкий слой, называемый *двойным электрическим слоем*, в котором существует сильное электрическое поле, напряженность которого направлена от воды к металлу. Это поле препятствует дальнейшему переходу ионов цинка в воду, и в результате возникает динамическое равновесие, при котором среднее число ионов, приходящих из металла в воду, равно числу ионов, возвращающихся из воды в металл.

Динамическое равновесие установится и в том случае, если металл погрузить в водный раствор соли того же металла, например цинк в раствор цинкового купороса $ZnSO_4$. В растворе соль диссоциирует на ионы Zn^{++} и SO_4^{--}. Образовавшиеся при этом ионы цинка ничем не отличаются от ионов цинка, поступивших в раствор с электрода. Повышение концентрации ионов цинка в электролите облегчает переход этих ионов в металл из раствора и затрудняет переход из металла в раствор. Поэтому в растворе цинкового купороса погруженный цинковый электрод хотя и заряжается отрицательно, но слабее, чем в чистой воде.

При погружении металла в раствор металл не всегда заряжается отрицательно. Например, если медный электрод погрузить в раствор медного купороса $CuSO_4$, то ионы Cu^{++} начнут из раствора осаждаться на электроде, заряжая его положительно. Напряженность поля в двойном электрическом слое в данном случае направлена от меди к раствору.

Таким образом, при погружении металла в воду или в водный раствор, содержащий ионы того же металла, на границе металла с раствором между ними возникает разность потенциалов. Знак и величина этой разности потенциалов зависит от типа металла (медь, цинк и т. д.), от концентрации ионов в растворе и почти не зависит от температуры и давления.

Два электрода из разных металлов, погруженные в электролит, образуют *гальванический элемент*. Например, в элементе Вольта цинковый и медный электроды погружены в водный раствор серной кислоты. В первый момент раствор не содержит ни ионов цинка, ни ионов меди. Однако в дальнейшем эти ионы поступают в раствор с электродов и устанавливается динамическое равновесие. Пока электроды не соединены друг с другом проводом, потенциал электролита одинаков во всех точках, а потенциалы электродов отличаются от потенциала электролита благодаря образующимся двойным слоям на их границе с электролитом. При этом электродный потенциал цинка равен $-0{,}763$ В, а меди $+0{,}337$ В. Электродвижущая сила элемента Вольта, складывающаяся из этих скачков потенциалов, будет равна $\mathscr{E} = 0{,}337 - (-0{,}763) = 1{,}10$ В.

Ток в цепи с гальваническим элементом. Если электроды гальванического элемента соединить проводом, то электроны по этому проводу будут переходить с отрицательного электрода (цинк) на положительный (медь), что нарушает динамическое равновесие между электродами и электролитом, в который они погружены. Ионы цинка начнут переходить с электрода в раствор, так чтобы поддерживать двойной электрический слой в прежнем состоянии с неизменным скачком потенциала между электродом и электролитом. Аналогично у медного электрода ионы меди начнут переходить из раствора и осаждаться на электроде. При этом около отрицательного электрода образуется недостаток ионов SO_4^{--}, а у положительного — избыток таких ионов. Общее число ионов SO_4^{--} в растворе не изменится.

В результате описанных процессов в замкнутой цепи будет поддерживаться электрический ток, который в соединительном проводе создается движением электронов, а в электролите ионами. При прохождении электрического тока происходит постепенное растворение цинкового электрода и осаждение меди на положительном (медном) электроде. Концентрация ионов SO_4^{--} увеличивается у цинкового электрода и уменьшается у медного.

Потенциал в цепи с гальваническим элементом. Описанная картина прохождения электрического тока в неоднородной замкнутой цепи, содержащей химический элемент, соответствует распределению потенциала вдоль цепи, схематически показанному на рис. 107. Во внешней цепи, т. е. в соединяющем электроды проводе потенциал плавно понижается от значения φ_1 на положительном (медном) электроде A до значения φ_2 на отрицательном (цинковом) электроде B в соответствии с законом Ома для однородного проводника. Во внутренней цепи, т. е. в электролите между электродами, потенциал плавно понижается от значения φ_2' вблизи цинкового электрода до значения φ_1' вблизи медного электрода. Если во внешней цепи ток идет от медного электрода к цинковому, то внутри электролита — от цинкового к медному. Скачки потенциалов $\varphi_1 - \varphi_1'$ и $\varphi_2 - \varphi_2'$ в двойных электрических слоях создаются в результате действия сторонних (в данном случае химических) сил. Движение электрических зарядов в двойных слоях благодаря сторонним силам происходит против направления действия электрических сил.

Наклонным участкам изменения потенциала на рис. 107 соответствуют электрические сопротивления внешнего и внутреннего участков замкнутой цепи. Суммарное падение потенциала вдоль этих участков равно сумме скачков потенциала в двойных слоях, т. е. электродвижущей силе (ЭДС) элемента.

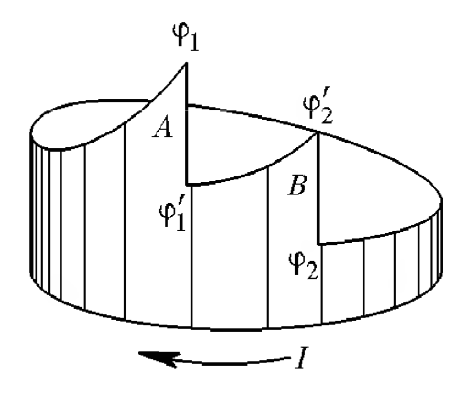


Рис. 107. Распределение потенциала вдоль цепи, содержащей химический элемент

Прохождение электрического тока в гальваническом элементе осложняется побочными продуктами, выделяющимися на электродах, и появлением перепада концентрации в электролите. Об этих явлениях говорят как об *электролитической поляризации*. Например, в элементах Вольты при замыкании цепи положительные ионы H^+ движутся к медному электроду и осаждаются на нем. В результате через некоторое время медный электрод как бы заменяется водородным. Так как электродный потенциал водорода на 0,337 В ниже электродного потенциала меди, то ЭДС элемента уменьшается примерно на такую же величину. Кроме того, выделяющийся на медном электроде водород увеличивает внутреннее сопротивление элемента.

Для уменьшения вредного влияния водорода используются деполяризаторы — различные окислители. Например, в наиболее употребительном элементе Лекланше («сухие» батарейки) положительным электродом служит графитовый стержень, окруженный спрессованной массой перекиси марганца MnO_2 и графита.

Аккумуляторы. Практически важной разновидностью гальванических элементов являются *аккумуляторы*, для которых после разрядки возможен обратный процесс зарядки с преобразованием электрической энергии в химическую. Вещества, расходуемые при получении электрического тока, восстанавливаются внутри аккумулятора путем электролиза.

Наибольшее распространение получили свинцовые, или кислотные, аккумуляторы. Электроды аккумулятора изготавливаются из двух решетчатых свинцовых пластин, заполненных пастой из окиси свинца PbO. В аккумулятор заливают 30%-ный раствор серной кислоты (с плотностью 1,2 г/см^3). На поверхностях пластин образуется слой сернокислого свинца $PbSO_4$. Чтобы аккумулятор стал источником тока, его необходимо зарядить. При пропускании постоянного электрического тока на электроде, соединенном с положительным полюсом выпрямителя, отрицательные ионы SO_4^{--} нейтрализуются и при соединении с сернокислым свинцом и водой образуют перекись свинца PbO_2 в соответствии с реакцией

$$PbSO_4 + SO_4^{--} + 2H_2O = PbO_2 + 2H_2SO_4 + 2e.$$

Положительные ионы H^+ движутся к электроду, соединенному с отрицательным полюсом выпрямителя, и нейтрализуются на нем в результате реакции

$$PbSO_4 + 2H^+ + 2e = Pb + H_2SO_4.$$

Видно, что при зарядке аккумулятора повышается концентрация серной кислоты, что ведет к увеличению плотности электролита.

Таким образом, в процессе зарядки создается резкая асимметрия электродов: один становится свинцовым, другой — из перекиси свинца. Заряженный аккумулятор представляет собой гальванический элемент, способный служить источником тока.

При подключении к аккумулятору потребителей электрической энергии через цепь потечет электрический ток, направление которого противоположно зарядному току. Химические реакции идут в обратном направлении и аккумулятор возвращается в исходнос состоянис. Оба элсктрода будут покрыты слосм соли $PbSO_4$, а концентрация серной кислоты вернется к первоначальному значению.

У заряженного аккумулятора ЭДС составляет примерно 2,2 В. При разрядке она понижается до 1,85 В. Дальнейшую разрядку производить не рекомендуется, так как процесс образования сернокислого свинца становится необратимым и аккумулятор портится.

Максимальный заряд, который может отдать аккумулятор при разрядке, называется его емкостью. Емкость аккумулятора обычно измеряется в ампер-часах. Она тем больше, чем больше поверхность пластин.

Применения электролиза. Электролиз используется в металлургии. Наиболее распространено электролитическое получение алюминия и чистой меди. С помощью электролиза можно создавать тонкие слои одних веществ на поверхности других с целью получения декоративных и защитных покрытий (никелирование, хромирование). Процесс получения отслаиваемых покрытий (гальванопластика) был разработан русским ученым Б. С. Якоби, применившим его для изготовления полых скульптур, украшающих Исаакиевский собор в Санкт-Петербурге.

- Чем отличается физический механизм электропроводности в металлах и электролитах?
- Поясните, почему степень диссоциации данного вещества зависит от диэлектрической проницаемости растворителя.
- Объясните, почему в сильно разбавленных растворах электролита практически все молекулы растворенного вещества диссоциированы.
- Поясните, в чем механизм электропроводности электролитов сходен с механизмом электропроводности газов. Почему при неизменных внешних условиях электрический ток пропорционален приложенному напряжению?
- Какую роль при выводе закона электролиза (3) играет закон сохранения электрического заряда?
- Поясните связь электрохимического эквивалента вещества с удельным зарядом его ионов.
- Как можно на опыте определить отношение электрохимических эквивалентов разных веществ, если имеется несколько электролитических ванн, но нет приборов для измерения силы тока?
- Каким образом явление электролиза можно использовать для создания счетчика расхода электроэнергии в сети постоянного тока?
- Почему законы Фарадея можно рассматривать как экспериментальное доказательство представлений об атомарной природе электричества?
- Какие процессы происходят при погружении металлических электродов в воду и в электролит, содержащий ионы этих металлов?
- Опишите процессы, происходящие в электролите вблизи электродов гальванического элемента при прохождении тока.
- Почему внутри гальванического элемента положительные ионы движутся от отрицательного (цинкового) электрода к положительному (медному) электроду? Каким образом в цепи возникает распределение потенциала, заставляющее ионы двигаться именно так?
- Почему степень заряженности кислотного аккумулятора можно проверять при помощи ареометра, т. е. прибора для измерения плотности жидкости?
- Чем в принципиальном отношении отличаются процессы в аккумуляторах от процессов в «сухих» батарейках?

- Какая часть электрической энергии, затраченной в процессе зарядки аккумулятора с ЭДС $\mathscr{E}$, может быть использована при его разрядке, если в процессе зарядки аккумулятора на его клеммах поддерживалось напряжение U?

§ 33. Плазма

Во всех рассмотренных выше явлениях, связанных с прохождением электрического тока в газах и жидкостях, взаимодействие носителей заряда друг с другом было несущественно из-за малой их концентрации. Основную роль играло взаимодействие электронов или ионов с нейтральными частицами — атомами и молекулами. Однако существует целый ряд явлений, в которых взаимодействие заряженных частиц друг с другом играет определяющую роль. В таком случае говорят о проявлении плазменных свойств. Что же такое плазма?

Вообще плазмой называют систему, состоящую из большого числа подвижных частиц, по крайней мере часть которых обладает электрическим зарядом. Термин «плазма» широко используется в современной физике. Этот термин применяют к ионизованному газу при таких условиях, когда силами взаимодействия составляющих его частиц нельзя пренебрегать. Плазмой называют и электронный газ в металлах и полупроводниках.

Плазма — это наиболее распространенное состояние вещества в природе. Звезды представляют собой гигантские сгустки горячей плазмы. Внешний ионизированный слой земной атмосферы, радиационные пояса, некоторые типы кометных хвостов, наконец пламя мартеновской печи — все это примеры систем, являющихся плазмой.

Коллективное движение частиц в плазме. Отличительные свойства плазмы связаны с тем, что в ее состав входят свободные частицы, обладающие электрическим зарядом. Новые, необычные для нейтрального газа свойства плазмы обусловлены тем чрезвычайно сильным воздействием, которое оказывают электрические и магнитные поля на движение заряженных частиц. На нейтральные частицы электрические и магнитные поля оказывают гораздо меньшее воздействие. В газе нейтральных частиц информация о локальном изменении состояния, например об увеличении концентрации частиц в каком-либо месте, передается только в результате столкновений частиц.

В плазме картина иная. В отличие от нейтральных частиц, которые взаимодействуют друг с другом только на малых расстояниях, заряженные частицы взаимодействуют с помощью дальнодействующих кулоновских сил. При локальном изменении состояния в плазме возникают электрическое и магнитное поля, которые действуют на всю плазму в целом. В результате в плазме возникают коллективные движения частиц — колебания и волны.

Скорость передачи информации о локальных возмущениях определяется скоростью распространения электромагнитных волн в плазме. Именно наличие таких специфических коллективных процессов в плазме — волн, в которых происходят колебания как частиц плазмы, так и сопровождающих движение заряженных частиц электромагнитных полей, — и позволяет говорить о плазме как о четвертом агрегатном состоянии вещества.

Газоразрядная плазма, как правило, представляет собой смесь трех компонент: свободных электронов, положительных ионов и нейтральных атомов или молекул. Электроны — это наиболее подвижная часть такой плазмы, и именно с движением электронов связаны ее наиболее интересные свойства. Ионы же вследствие гораздо большей массы ведут себя более «пассивно», благодаря чему во многих случаях можно вообще пренебрегать их движением и рассматривать ионную часть плазмы как неподвижный положительный фон, на котором происходит движение электронов. Такое приближение однокомпонентной плазмы тем более оправдано для электронов в металлах, где движение ионов вообще ограничено колебаниями вблизи положения равновесия.

Квазинейтральность плазмы. Одним из важнейших свойств плазмы является ее стремление к сохранению равенства плотностей электрического заряда положительных и отрицательных частиц. В самом деле, при высокой плотности заряженных частиц в плазме даже малое пространственное разделение положительных и отрицательных зарядов привело бы к появлению очень сильных электрических полей, стремящихся восстановить локальное нарушение электронейтральности. Поэтому в среднем (в достаточно большом объеме или за достаточно большой промежуток времени) плазма должна быть почти нейтральной, или, как говорят, *квазинейтральной*.

Плазменные колебания. Оценим размеры объема и промежуток времени, в которых выполняется квазинейтральность плазмы. Представим себе, что в плоском слое однородной нейтральной в целом плазмы все электроны сместились на расстояние x в одном и том же направлении (рис. 108). Возникающее в результате такого смещения электронов результирующее распределение зарядов будет таким же, как в плоском конденсаторе. Электрическое поле в плазме определяется плотностью заряда на «обкладках» такого «конденсатора». При смещении электронов в слое толщины l (рис. 108)

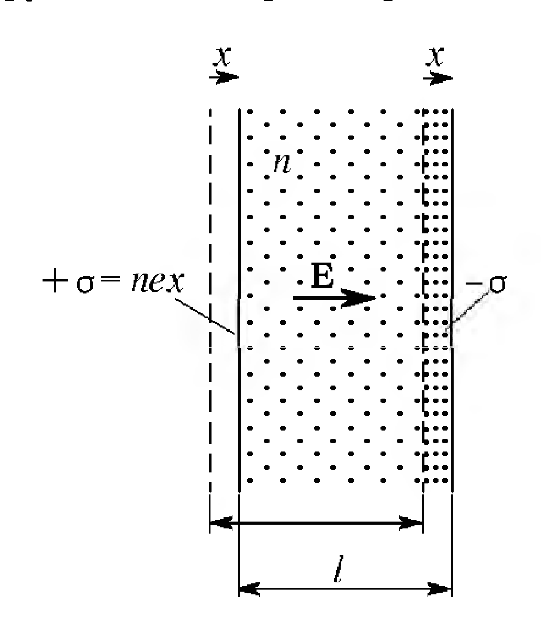


Рис. 108. К выводу формулы для частоты плазменных колебаний

нарушение нейтральности происходит только в тонких областях толщины x вблизи границ слоя: слева образуется избыток положительного заряда, справа — отрицательного. Если концентрацию электронов в нейтральной плазме обозначить через n, то при смещении всех электронов в слое на расстояние x заряд σ на единице площади «обкладки» будет равен enx. Поэтому напряженность поля E в конденсаторе будет равна (в гауссовой системе единиц)

$$E = -4\pi\sigma = 4\pi enx. \tag{1}$$

Действующая на каждый электрон в слое сила $\mathbf{F} = e\mathbf{E}$, будет пропорциональна смещению электрона x и, как видно из рис. 108, направлена в сторону, противоположную смещению:

$$F = -4\pi ne^2x. \tag{2}$$

Поэтому электроны будут совершать гармонические колебания, частота которых ω_p определяется выражением

$$\omega_p^2 = \frac{4\pi ne^2}{m}, \tag{3}$$

где m — масса электрона. В СИ выражение для частоты ω_p имеет вид

$$\omega_p^2 = \frac{ne^2}{\varepsilon_0 m}. \tag{3a}$$

Частота ω_p — одна из важнейших характеристик плазмы. Ее называют *плазменной частотой*, а сами колебания плазменными или ленгмюровскими, по имени американского физика И. Ленгмюра, впервые исследовавшего эти колебания.

Таким образом, в результате разделения зарядов в плазме возникают электрические поля, вызывающие колебания частиц. Эти колебания стремятся восстановить квазинейтральность плазмы. Ясно, что заметить отклонения плазмы от квазинейтральности можно только на протяжении времени, малого по сравнению с периодом плазменных колебаний. В среднем (за много периодов колебаний) плазма ведет себя как квазинейтральная среда. Период плазменных колебаний $T_p = 2\pi/\omega_p$ — это характерный временно́й масштаб разделения зарядов в плазме.

Экранировка кулоновского взаимодействия. Теперь оценим пространственный масштаб разделения зарядов в классической плазме. Чем может быть вызвано самопроизвольное разделение зарядов? Очевидно, флуктуациями пространственного распределения частиц, обусловленными их хаотическим тепловым движением. Характерная скорость v теплового движения электронов зависит от термодинамической температуры T: $\langle v\rangle \sim \sqrt{kT/m}$. Пространственный масштаб разделения зарядов определяется произведением характерной тепловой скорости на время движения в одном направлении. Чем же определяется это время?

В отличие от газа из нейтральных частиц, где это время равно времени свободного движения электрона между столкновениями с атомами или молекулами, в плазме характерное время свободного движения электрона определяется периодом плазменных колебаний $T_p = 2\pi/\omega_p$. Действительно, при пространственном разделении зарядов возникают плазменные колебания, которые стремятся восстановить электронейтральность. При колебаниях скорость электрона за половину периода изменяет свое направление на противоположное.

Таким образом, для оценки размера области пространственного разделения зарядов нужно характерную тепловую скорость $\sqrt{kT/m}$ умножить на время движения, которое по порядку величины есть $1/\omega_p = \sqrt{m/4\pi ne^2}$:

$$r_D \sim \sqrt{\frac{kT}{m}}\,\frac{1}{\omega_p} = \sqrt{\frac{kT}{4\pi ne^2}}\,. \qquad (4)$$

Эта величина носит название дебаевского радиуса экранирования, по имени английского физика П. Дебая, который впервые ввел ее при изучении экранировки кулоновского взаимодействия заряженных частиц в электролитах.

На расстоянии дебаевского радиуса происходит экранирование кулоновского поля любого заряда. Причиной этого экранирования является преимущественная группировка вокруг любого заряда заряженных частиц противоположного знака. Кулоновские силы стремятся максимально приблизить к внесенному в плазму пробному заряду заряженные частицы противоположного знака, а хаотическое тепловое движение препятствует этому. В результате вокруг пробного заряда возникает пространственно неоднородное распределение электронов, приводящее к экранировке его кулоновского поля на расстоянии порядка дебаевского радиуса r_D.

Пространственные масштабы в плазме. Чтобы плазма сохраняла квазинейтральность, занимаемая ею область (т. е. ее линейный размер L) должна намного превосходить дебаевский радиус r_D: $r_D \ll L$. Только при выполнении условия квазинейтральности плазма проявляет коллективные свойства, а не ведет себя как простая совокупность заряженных частиц. Из рассмотренной выше физической картины экранирования кулоновского взаимодействия следует, что дебаевский радиус экранирования должен намного превосходить среднее расстояние $\langle r\rangle$ между заряженными частицами: $\langle r\rangle \ll r_D$.

Итак, чтобы электронейтральная в целом система электронов и ионов могла рассматриваться как плазма, необходимо выполнение условий

$$\langle r\rangle \ll r_D \ll L. \qquad (5)$$

Первому из условий (5) можно придать несколько иной вид. Так как среднее расстояние $\langle r\rangle$ определяется концентрацией n электро-

нов ($\langle r \rangle \sim n^{-1/3}$), то с помощью (4) для r_D это условие переписывается следующим образом:

$$\frac{e^2 n^{1/3}}{kT} \ll 1. \qquad (6)$$

В числителе этого выражения стоит энергия кулоновского взаимодействия частиц, находящихся на расстоянии $\langle r \rangle$ друг от друга, а в знаменателе — характерная энергия их хаотического теплового движения. Таким образом, в классической плазме (т. е. когда движение частиц подчиняется законам классической физики) средняя энергия кулоновского взаимодействия мала по сравнению с энергией теплового движения.

Пространственная однородность плазмы. Как совместить представления о пространственной неоднородности распределения зарядов в плазме при экранировании поля отдельных зарядов с макроскопической однородностью плазмы в целом?

Чтобы экспериментально обнаружить пространственную неоднородность распределения зарядов, нужно ввести в плазму зонды для измерения локальных характеристик электрического поля. Если такие измерения возможны для областей, размеры которых меньше дебаевского радиуса, а времена измерения меньше периода плазменных колебаний, то локальные изменяющиеся со временем неоднородности действительно будут обнаружены. Для реальных приборов, обладающих инерционностью, локальные поля за время наблюдения усредняются и в равновесии плазма предстает как стационарная пространственно однородная система.

Плазма и анализ размерностей. Характерный для плазмы пространственный масштаб (дебаевский радиус r_D) и характерный промежуток времени (период плазменных колебаний) мы оценили на основе наглядной физической картины движения зарядов в плазме. Однако это можно сделать из соображений размерности, не конкретизируя картины движения частиц в рассматриваемой модели плазмы. Покажем это.

Полностью ионизованная однокомпонентная квазинейтральная плазма в состоянии термодинамического равновесия характеризуется четырьмя параметрами: зарядом частиц e, их массой m, концентрацией n и температурой T. Нетрудно убедиться, что из этих величин можно составить только один независимый безразмерный параметр γ:

$$\gamma = \frac{e^2 n^{1/3}}{kT}, \qquad (7)$$

совпадающий с левой частью неравенства (6). Постоянная Больцмана k в этом выражении появляется потому, что в системе СГС температуру, характеризующую среднюю энергию теплового движения, необходимо выразить в эргах. Другими словами, если через T обозначена температура в кельвинах, то в эргах ей соответствует величина kT.

Для того чтобы найти характерный для рассматриваемой системы параметр, имеющий размерность времени, следует сначала найти любой параметр, имеющий такую размерность. Можно взять, например, отношение среднего расстояния между частицами $\langle r\rangle \sim n^{-1/3}$ к скорости теплового движения $\langle v\rangle \sim \sqrt{kT/m}$. Далее этот параметр следует умножить на произвольную функцию безразмерного параметра, существующего для этой системы.

Итак, самый общий вид параметра, имеющего размерность времени,

$$t = n^{-1/3}\sqrt{m/kT}\,f(\gamma). \tag{8}$$

Формула (8) даже без конкретизации вида функции $f(\gamma)$ позволяет сделать интересные выводы. Прежде всего отметим, что в рассмотренной выше модели газа из нейтральных частиц, где $\gamma = nr_0^3$ (r_0 — радиус межмолекулярного взаимодействия), отсутствует параметр, имеющий размерность времени, не зависящий от температуры газа. Другими словами, все возможные в такой системе неравновесные процессы определяются скоростями хаотического теплового движения молекул.

Иначе обстоит дело в плазме, где безразмерный параметр γ, даваемый формулой (7), сам содержит температуру T. Легко видеть, что если в (8) выбрать в качестве произвольной функции $f(\gamma) = \gamma^{-1/2}$, то для параметра t, имеющего размерность времени, получится выражение, не содержащее термодинамической температуры T, а тем самым и скорости теплового движения частиц:

$$t = n^{-1/3}\left(\frac{m}{kT}\right)^{1/2}\left(\frac{kT}{n^{1/3}e^2}\right)^{1/2} = \left(\frac{m}{ne^2}\right)^{1/2}. \tag{9}$$

Это значение, как видно из формулы (3) для плазменной частоты ω_p, совпадает с периодом плазменных колебаний $2\pi/\omega_p$.

Независимость характерного времени от скорости — это признак гармонических колебаний. Таким образом, только из анализа размерностей мы также приходим к идее плазменных колебаний, качественная физическая картина возникновения которых была рассмотрена выше. Продолжая аналогичные рассуждения, можно прийти к понятию и к оценке значения дебаевского радиуса экранирования r_D.

Приведем числовые значения характерных ленгмюровских частот и дебаевских радиусов. Для типичной классической лабораторной плазмы с концентрацией электронов $n \approx 10^{14}$ см$^{-3}$ получаем $\omega_p \approx 5\cdot 10^{11}$ с$^{-1}$, что примерно в тысячу раз меньше частоты видимого света. Значение дебаевского радиуса r_D зависит от температуры. В частности, при термоядерных температурах ($kT \approx 10^4$ эВ) $r_D \approx 10^{-2}$ см.

Рассматривая плазму как колебательную систему с характерной собственной частотой ω_p, естественно ввести представление о диск-

ретных квантовых энергетических состояниях. Квант энергии колебаний с частотой ω_p согласно формуле Планка $E = \hbar\omega$ для приведенного выше значения ω_p составляет $5 \cdot 10^{-16}$ эрг $\approx 3 \cdot 10^{-4}$ эВ. Уже при комнатной температуре $T \approx 300$ K энергия теплового движения частиц равна $kT \approx 4 \cdot 10^{-14}$ эрг $\approx 1/40$ эВ. Это означает, что квантование плазменных уровней энергии не играет роли и тепловое движение частиц приводит к возбуждению плазменных колебаний, что на квантовом языке соответствует рождению большого числа квантов колебаний — плазмонов.

Волны в плазме. Выведенная из состояния термодинамического равновесия плазма стремится наиболее быстрым путем релаксировать к равновесному состоянию. В отличие от газа нейтральных частиц, где такой процесс релаксации определяется исключительно передачей энергии при межчастичных столкновениях, плазма избавляется от неравновесности преимущественно путем возбуждения колебаний и волн. Это так называемые плазменные неустойчивости, которые оказались основной преградой на пути осуществления управляемой термоядерной реакции.

Плазменные колебания — это по существу электрические продольные колебания, в которых магнитное поле практически отсутствует, а вектор напряженности электрического поля направлен вдоль движения частиц. Кроме таких продольных колебаний, в плазме, как и в других материальных средах, возможно распространение обычных поперечных электромагнитных волн. Для них плазменная частота ω_p представляет собой граничную частоту. Электромагнитные волны с частотой меньше ω_p не могут проникать в плазму, так как их низкочастотное электрическое поле экранируется свободными электронами плазмы. Падающая на границу плазмы волна при $\omega < \omega_p$ отражается от границы. Если же частота электромагнитной волны выше плазменной частоты, то такая волна проникает в плазму.

Для ионосферной плазмы граница прозрачности попадает в диапазон метровых радиоволн. Поэтому радиосвязь с объектами в космосе возможна только в диапазоне метровых и дециметровых волн. Отражение волн длиной в несколько десятков метров от ионосферы используется для земной радиосвязи на большие расстояния.

Применения плазмы. Плазма широко применяется в технике. Газоразрядная плазма служит активной средой в газовых лазерах — квантовых источниках когерентного оптического излучения. Плазменные струи используются в новых перспективных магнитогидродинамических генераторах электрической энергии. В новых приборах — плазмотронах — создаются мощные струи плотной плазмы, применяемые для резки и сварки металлов, бурения скважин в твердых породах и т. д. В плазме ускоряются многие химические ре-

акции и могут происходить реакции, невозможные в обычных условиях. Большие надежды связаны с обузданием высокотемпературной плазмы для осуществления управляемой реакции термоядерного синтеза.

- Чем отличается поведение носителей заряда в плазме и в газе из нейтральных атомов?
- Что такое квазинейтральность плазмы? Чем квазинейтральность отличается от истинной нейтральности?
- Какова причина существования плазменных ленгмюровских колебаний? Чем определяется их частота?
- Как можно оценить радиус экранирования электрического поля внесенного в плазму пробного заряда?
- В каком соотношении находятся дебаевский радиус экранирования со средним расстоянием между электронами плазмы и размерами занимаемой плазмой области?
- Как совместить представления о пространственной однородности плазмы с экранированием кулоновского взаимодействия зарядов плазмы?
- Поясните, почему в плазме, в отличие от нейтрального газа, существует характерный временной масштаб, не зависящий от температуры.
- Чем плазменные колебания отличаются от электромагнитных волн, которые могут распространяться в плазме?
- Что такое плазменные неустойчивости и в чем причина их возникновения?
- Какие практические применения плазмы вам известны?

VII. ЭЛЕКТРОННЫЕ СВОЙСТВА ТВЕРДЫХ ТЕЛ

§ 34. Электронная структура кристаллов

Любое макроскопическое тело состоит из атомов и молекул, подчиняющихся законам квантовой физики. Поэтому объяснение большинства наблюдаемых свойств макроскопических тел на основе представлений об их микроструктуре невозможно без использования квантовых законов.

Диэлектрики, полупроводники, металлы. Наибольший прогресс в применении квантовой механики к макроскопическим системам достигнут при изучении твердых тел, т. е. тел, обладающих кристаллической структурой. Первое, что нужно знать для объяснения наблюдаемых электрических, магнитных, тепловых, оптических и других макроскопических свойств, — это уровни энергии электронов в кристалле. Строго говоря, в кристалле, как и в отдельном атоме, можно рассматривать только состояние всей системы в целом. Тем не менее, как и в атоме, с хорошей точностью можно говорить о состояниях отдельных электронов в некотором потенциальном пространственно-периодическом поле кристалла.

Уровни энергии электронов в кристаллах. Получить качественное представление о структуре энергетического спектра электронов в твердом теле можно, проследив за тем, как уровни энергии изолированных атомов изменяются при объединении этих атомов в кристалл. Допустим, что N одинаковых атомов расположены в пространственной решетке со столь большим межатомным расстоянием, что их взаимодействием друг с другом можно пренебречь. Ясно, что энергетические уровни электронов в таком гипотетическом кристалле будут такие же, как и у изолированного атома. Разница будет только в том, что теперь каждому уровню энергии соответствует в N раз больше различных электронных состояний, чем в одном атоме.

Будем постепенно уменьшать межатомное расстояние. По мере сближения все более и более существенным становится взаимодействие между атомами, которое сказывается на уровнях энергии электронов. Равновесное расстояние между атомами в кристалле приблизительно таково, что электронные оболочки, соответствующие внешним (валентным) электронам, приходят в соприкоснове-

ние. При этом электроны внутренних электронных оболочек, размеры которых малы по сравнению с межатомным расстоянием в кристалле, почти не «чувствуют» поля, создаваемого соседними атомами, и их состояние в кристалле почти не отличается от состояния в изолированном атоме.

Энергетические зоны. Наиболее сильному возмущающему воздействию соседних атомов подвергаются самые удаленные от ядра электроны. В результате взаимодействия с соседними атомами каждый уровень энергии валентных электронов расщепляется на большое число близко расположенных уровней, которые можно рассматривать как квазинепрерывную зону разрешенных значений энергии электронов.

Так как состояния исходной системы удаленных друг от друга атомов изменяются при их сближении непрерывным образом, то число разрешенных квантовых состояний электронов в кристалле должно быть таким же, как в исходной системе. Поэтому число уровней в каждой зоне равно полному числу атомов в кристалле N. Однако ширина разрешенной энергетической зоны не зависит от полного числа атомов в кристалле, а для данного вещества определяется только межатомным расстоянием. При увеличении числа атомов в кристалле возрастает лишь густота энергетических уровней в пределах разрешенной зоны. Так как число атомов в кристалле велико ($N \sim 10^{23}$), то энергетический спектр электронов в пределах разрешенной зоны можно считать практически непрерывным.

Нас интересуют главным образом энергетические зоны, соответствующие внешним, т. е. валентным, электронам атомов, ибо именно они определяют большинство наблюдаемых макроскопических свойств кристаллов. Если уровень энергии валентного электрона атома в основном состоянии был заполнен электронами целиком, т. е. были заняты все различные состояния с данным значением энергии, то и в кристалле соответствующая энергетическая зона будет заполнена полностью. В этом случае зона называется *валентной*.

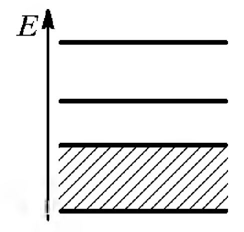


Рис. 109. Схема энергетических зон кристалла

Валентная зона отделена от расположенной выше разрешенной зоны, возникшей из наинизшего возбужденного уровня энергии атома, некоторым энергетическим интервалом, называемым *запрещенной зоной*. Когда кристалл находится в основном состоянии, т. е. состоянии с наинизшей энергией, все разрешенные зоны, расположенные выше валентной, пусты — в соответствующих этим зонам состояниях электронов нет. Диаграмма энергетических зон кристалла схематически показана на рис. 109. Заполненные электронами уровни энергии заштрихованы.

Заполнение зон в диэлектриках. Такое заполнение энергетических зон электронами характерно для *диэлектриков.* В полностью заполненной валентной зоне диэлектриков свободных уровней энергии нет. Поэтому под действием приложенного к кристаллу электрического поля электрон в заполненной зоне не может изменить своего состояния. Этим объясняется отсутствие электропроводности у диэлектриков.

Собственные полупроводники. Если в кристалле ширина запрещенной зоны (т. е. расстояние от заполненной валентной зоны до ближайшей свободной зоны) невелика, так что в тепловом равновесии при конечной температуре часть электронов из валентной зоны в результате теплового возбуждения оказывается в свободной зоне, то кристалл представляет собой *собственный полупроводник.* Находящиеся в почти пустой зоне электроны под действием внешнего электрического поля могут изменить свое состояние, т. е. ускоряться. Это означает, что в таком кристалле электрическое поле создает ток. Поэтому находящиеся в почти пустой зоне электроны называют *электронами проводимости*, а саму зону -- *зоной проводимости.* Так как с ростом температуры число электронов проводимости увеличивается, электропроводность полупроводников растет с температурой.

Примесные полупроводники. Кроме собственных полупроводников, в которых электрический ток возможен благодаря тому, что часть электронов из валентной зоны тепловое движение переводит в зону проводимости, существуют так называемые *примесные полупроводники.* В примесных полупроводниках носители тока появляются благодаря переходам электронов не из одной энергетической зоны в другую, а с энергетических уровней атомов примеси в свободную энергетическую зону, или из заполненной зоны на свободные энергетические уровни примесей. В этом случае электрические свойства кристалла очень сильно зависят от концентрации примесей.

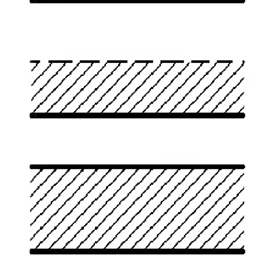


Рис. 110. В металле зона проводимости заполнена электронами частично

Металлы. Если в основном состоянии изолированного атома уровень энергии валентного электрона был заполнен частично, т. е. только часть разрешенных состояний с данной энергией занята электронами, то и при образовании кристалла из таких атомов соответствующая энергетическая зона будет заполнена лишь частично (рис. 110). В этом случае в зоне имеются нс занятыс элсктронами состояния, т. с. возможсн элсктричсский ток под действием приложенного поля. Так как число электронов в этой частично заполненной зоне проводимости очень велико не меньше, чем число атомов в кристалле, — то и проводимость такого кристалла велика. Это — *металлы.*

Характерное для металлов заполнение зон может получиться и в том случае, когда у изолированного атома уровень энергии валентных электронов заполнен целиком, но при сближении атомов в кристалл происходит настолько сильное расщепление уровней, что верхняя целиком заполненная зона и соседняя с ней пустая зона начинают перекрываться. В этом случае наименьшее значение энергии кристалла соответствует такому заполнению энергетических зон, когда часть электронов с верхних уровней заполненной зоны переходит на нижние уровни пустой зоны (рис. 111).

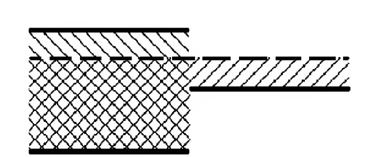


Рис. 111. Распределение электронов по энергетическим уровням в случае перекрытия свободной и заполненной зон

Таким образом, деление твердых тел на диэлектрики, полупроводники и металлы, основанное в первую очередь на различии их электрического сопротивления, имеет под собой глубокую основу, связанную с различием их электронной структуры. Различный характер заполнения энергетических зон в диэлектриках и металлах приводит к исключительно большому различию их электрических сопротивлений. Так, у чистых металлов при низких температурах удельное сопротивление может быть всего лишь 10^{-10} Ом·м, в то время как у диэлектриков с достаточно чистой поверхностью оно может достигнуть 10^{24} Ом·м. Таким образом, сопротивление диэлектриков и металлов может отличаться в 10^{30} раз!

Различие между проводниками и диэлектриками проявляется не только в электрическом сопротивлении, но и во многих других свойствах, например оптических: чистые однородные диэлектрические кристаллы прозрачны в видимой области спектра, а металлы — нет. Зато они хорошо отражают видимый свет.

Стационарные состояния электронов в кристалле. Рассмотрим теперь подробнее стационарные состояния электронов, соответствующие какой-либо разрешенной энергетической зоне. Поскольку входящие в состав твердого тела атомы образуют правильную кристаллическую решетку, то потенциальное поле, действующее на какой-либо электрон со стороны ядер и всех остальных электронов, имеет пространственно-периодический характер.

Квантовомеханическое решение задачи о движении электрона в поле периодического потенциала приводит к следующим результатам. Стационарные состояния электрона в таком поле во многом напоминают состояния свободного электрона. Состояние свободной частицы характеризуется определенным значением импульса $\mathbf{p}$, поскольку для свободной частицы импульс является сохраняющейся величиной. Так как импульс имеет строго определенное значение, то вследствие соотношения неопределенностей Гейзенберга координаты электрона не имеют определенного значения: в таком состоя-

нии электрон как бы «размазан» по всему пространству в том смысле, что вероятность обнаружить его в любом месте одинакова.

Квазиимпульс. Как и у свободной частицы, состояния электрона в периодическом поле характеризуются вектором **p**. В отличие от импульса свободной частицы вектор **p** в кристалле изменяется в некоторой ограниченной области, размер которой зависит от расстояния между атомами.

Так можно считать потому, что энергия электрона $\varepsilon(\mathbf{p})$ в кристалле представляет собой периодическую функцию импульса. Состояния электрона с одинаковой энергией, в которых импульс **p** отличается на период функции $\varepsilon(\mathbf{p})$, физически эквивалентны. Это означает, что область изменения импульса электрона в кристалле можно ограничить ячейкой с центром в точке $\mathbf{p} = 0$ пространства импульсов. Такая ячейка называется *зоной Бриллюэна*. Подчеркнем, что зона Бриллюэна определена как ячейка пространства импульсов. Ее не следует путать с энергетической зоной, понимаемой как интервал разрешенных значений энергий.

Итак, значения импульсов электронов в кристалле, соответствующие всем возможным состояниям, лежат в зоне Бриллюэна, форма которой определяется типом симметрии кристалла. Оказывается, что размеры зоны Бриллюэна обратно пропорциональны расстояниям между атомами кристалла. Все это справедливо для электронов, находящихся в любой энергетической зоне. Чтобы подчеркнуть отличия импульса электрона в кристалле от импульса свободной частицы, который может принимать любые значения, лежащий в пределах зоны Бриллюэна импульс электрона в кристалле называют *квазиимпульсом*. Однако, как и свободная частица, электрон в кристалле, находящийся в стационарном состоянии с определенным значением квазиимпульса, не локализован, т. е. с равной вероятностью может быть обнаружен вблизи любого узла кристаллической решетки.

Эффективная масса. В каждой из разрешенных энергетических зон периодическая функция $\varepsilon(\mathbf{p})$, дающая зависимость энергии электрона от его квазиимпульса, имеет достаточно сложный вид. Однако в окрестностях точек пространства импульсов, соответствующих минимумам и максимумам $\varepsilon(\mathbf{p})$, ее можно аппроксимировать квадратичной функцией квазиимпульса. В частности, вблизи минимума энергии в некоторых кристаллах зависимость энергии электрона от **p** можно представить в таком же виде, как и для свободной частицы, с той только разницей, что масса свободного электрона должна быть заменена на некоторую *эффективную массу* m^*:

$$\varepsilon(\mathbf{p}) = \frac{p^2}{2m^*}, \qquad (1)$$

где квазиимпульс $\mathbf{p}$ отсчитывается от точки пространства импульсов, соответствующей минимуму энергии, а энергия отсчитывается от дна энергетической зоны.

В той области пространства импульсов, где энергия электрона достигает максимума, выражение для энергии электрона можно записать в виде

$$\varepsilon(\mathbf{p}) = \varepsilon_0 - \frac{p^2}{2m^*}, \tag{2}$$

где ε_0 — значение энергии в точке максимума. Значения эффективной массы m^*, фигурирующие в формулах (1) и (2), разные в разных энергетических зонах.

Эффективная масса m^* может быть как больше, так и меньше массы свободного электрона m. Более того, она может быть анизотропной, если свойства действующего на электрон периодического поля кристаллической решетки различны по разным направлениям. Понятие эффективной массы в физике твердого тела вводится для того, чтобы максимально приблизить описание движения электрона в разрешенной энергетической зоне к движению свободного электрона.

Зависимость энергии электронов зоны проводимости и валентной зоны от квазиимпульса $\mathbf{p}$ схематически показана на рис. 112 для случая, когда минимум зоны проводимости и максимум валентной зоны расположены в одной точке пространства импульсов (при $\mathbf{p} = 0$). В соответствии с формулами (1) и (2) графики этих зависимостей представляют собой параболы, причем для валентной зоны ветви такой параболы направлены вниз.

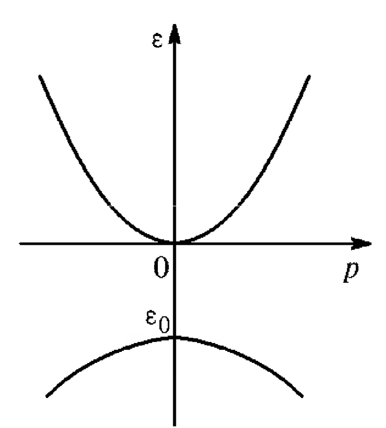


Рис. 112. Зависимость энергии электронов зоны проводимости и валентной зон от квазиимпульса

Электроны и дырки в полупроводниках. Концепция эффективной массы особенно полезна при описании электронных свойств полупроводников. Соотношение (1) используется для электронов в зоне проводимости, когда заполнены лишь состояния вблизи дна зоны. Соотношение (2) используется при описании движения электронов вблизи потолка заполненной валентной зоны, где при конечных температурах освобождаются состояния за счет перехода некоторого количества электронов в зону проводимости. Электрон, принадлежащий к такой группе энергетических состояний, реагирует на внешнее электрическое поле так, как если бы он обладал отрицательной массой, т. е. приобретал бы ускорение в направлении, противоположном действующей на него силе.

Движение всех электронов валентной зоны, в которой не заполнено лишь одно состояние вблизи ее потолка, эквивалентно движению некоторой частицы, которую обычно называют *положительной дыркой* или просто *дыркой*. Дырка имеет положительный заряд, так как соответствует отсутствию электрона в данном состоянии, и положительную массу m^*, так как под действием поля приобретает ускорение в направлении силы. Для зоны, которая заполнена почти до потолка, как в случае валентной зоны полупроводника (или для не полностью заполненных зон в некоторых металлах), более удобно говорить о движении дырок, а не электронов. Поэтому при описании электрического тока о вкладе зоны проводимости и валентной зоны говорят соответственно как об электронной и о дырочной проводимости.

Движение электронов и дырок в одной и той же энергетической зоне под действием приложенного электрического поля можно пояснить с помощью следующей грубой механической аналогии. Рассмотрим четыре стеклянные трубки, запаянные с обоих концов, которые расположены горизонтально. Пусть одна из них будет пустой, другая целиком заполнена водой, третья почти целиком заполнена водой, а четвертая содержит небольшое количество воды. Если один из концов каждой трубки приподнять, то на воду в трубках будет действовать сила, подобно тому, как внешнее электрическое поле действует на электроны в кристалле. В пустой трубке, очевидно, не произойдет никакого переноса. Также не произойдет никакого переноса в полностью заполненной трубке, поскольку частицы воды, подобно электронам в заполненной зоне, могут только обмениваться местами. Однако можно наблюдать, как вода перемещается от одного конца к другому в почти пустой трубке, в то время как пузырьки воздуха движутся в противоположном направлении в почти полной трубке. В последнем случае происходящие в трубке изменения удобно описывать на языке движения пузырьков воздуха, хотя внешние силы приводят в движение именно воду.

Об электропроводности кристаллов. Описанная выше картина движения электронов в кристаллах позволяет сделать некоторые выводы о механизме электропроводности. Проводимость может быть обусловлена электронами зоны, заполнение которой может меняться в широких пределах: от почти пустой до почти заполненной. Она, очевидно, невозможна в зоне, которая совершенно пуста. Она невозможна и в целиком заполненной зоне: если все состояния в энергетической зоне заняты электронами, то принцип Паули не допускает никакого отклика на внешнее электрическое поле. Электроны целиком заполненной зоны могут лишь обмениваться состояниями, но это ни к чему не приводит, поскольку электроны неразличимы.

Проводимость в почти пустой зоне осуществляется электронами с положительной массой и отрицательным зарядом. Проводимость почти заполненной зоны обусловлена движением небольшого числа незанятых состояний — положительных дырок, каждая из которых имеет положительную эффективную массу и положительный электрический заряд.

Равновесные и неравновесные носители. В собственных полупроводниках при отличных от абсолютного нуля температурах имеются как электроны в зоне проводимости, так и дырки в валентной зоне. Другими словами, при конечной температуре полупроводник находится не в основном состоянии (т. е. состоянии с наинизшей энергией), а в некотором возбужденном состоянии, которое, однако, является состоянием термодинамического равновесия с характерным для него распределением электронов по разрешенным уровням энергии.

Электроны и дырки, концентрация которых соответствует такому термодинамическому равновесию, называют *равновесными*. Сильной зависимостью равновесной концентрации свободных носителей от температуры объясняется быстрый рост электропроводности полупроводника при увеличении температуры.

Свободные носители тока в полупроводнике могут существовать не только благодаря термическому возбуждению кристалла, но и в результате внешних воздействий. Например, при освещении кристалла в результате поглощения фотонов могут происходить переходы электронов с уровней энергии валентной зоны на уровни энергии зоны проводимости (внутренний фотоэффект). Для этого энергия фотона должна быть не меньше ширины запрещенной зоны. Каждый поглощенный фотон рождает дополнительно пару электрон—дырка. В результате при освещении концентрация свободных носителей становится выше равновесной. Поэтому возбуждаемые светом носители тока называют *неравновесными*. Обусловленное светом увеличение электропроводности кристалла называют *фотопроводимостью*.

Наряду с процессами фотовозбуждения неравновесных носителей идет обратный процесс *рекомбинации* электронов и дырок, при котором электрон возвращается из зоны проводимости на свободное место в валентной зоне. Об этом процессе иногда говорят как об аннигиляции электронно-дырочной пары.

При продолжительном освещении устанавливается стационарное значение концентрации неравновесных носителей. После прекращения освещения концентрация свободных носителей постепенно возвращается к своему равновесному значению. При этом концентрация неравновесных носителей убывает со временем по экспоненциальному закону, а характерное время ее уменьшения (за которое концентрация убывает в $e \approx 2{,}72$ раза) называется *временем жизни* неравновесных носителей тока.

Как и свободная частица, электрон в кристалле под действием постоянного электрического поля движется равноускоренно. Иными словами, идеально регулярное периодическое потенциальное поле не оказывает сопротивления электрическому току. Закон Ома в идеальном кристалле не может выполняться, так как он имеет место только тогда, когда движение зарядов в электрическом поле происходит с постоянной средней скоростью. Сопротивление электрическому току в реальном кристалле обусловлено не потенциальными барьерами периодического поля, а отступлениями поля реального кристалла от строгой периодичности либо за счет тепловых колебаний атомов, образующих кристалл, либо за счет различного рода дефектов решетки — примесей, вакансий и т. п.

- Как объяснить существование разрешенных и запрещенных энергетических зон электронов в кристаллах исходя из представлений о дискретных энергетических уровнях электронов в изолированных атомах?
- Чем отличается заполнение энергетических зон в полупроводниках, диэлектриках и металлах?
- Чем отличаются друг от друга собственные и примесные полупроводники?
- Что такое квазиимпульс электрона в кристалле и чем он отличается от импульса свободной частицы?
- Что такое зона Бриллюэна? Объясните, зачем вводится это понятие.
- Почему в стационарном состоянии в кристалле, которое характеризуется определенным значением квазиимпульса, электрон не локализован, т. е. с равной вероятностью может быть обнаружен вблизи любого узла решетки?
- Что понимают под эффективной массой электрона в кристалле? Почему для зависимости энергии электрона от квазиимпульса используются разные выражения (1) и (2)?
- Что такое положительная дырка? В каких случаях движение электронов в кристалле удобно описывать, используя понятие дырки?
- Поясните роль принципа Паули в объяснении невозможности вклада электронов целиком заполненной энергетической зоны в электропроводность кристалла.
- Что такое равновесные и неравновесные носители тока в кристалле? Каким способом могут создаваться неравновесные носители?
- Может ли выполняться закон Ома в идеальной строго периодической кристаллической структуре? Почему?
- Какие причины вызывают сопротивление электрическому току в реальном кристалле?

§ 35. Электронные свойства металлов

Многие макроскопические свойства металлов, обусловленные наличием в них электронов проводимости, невозможно объяснить на основе классической статистической механики. Объяснение поведения

коллектива электронов проводимости требует обязательного привлечения квантовых закономерностей.

Принцип Паули и электроны в металлах. Рассмотрим подробнее свойства электронов проводимости в металлах, считая для простоты, что их эффективная масса изотропна, а число электронов проводимости в образце равно числу атомов металла N. Именно так обстоит дело, например, у щелочных металлов. Хотя поведение отдельного электрона проводимости в кристалле очень сходно с поведением классической свободной частицы, свойства коллектива таких электронов резко отличаются от свойств классического газа. Причина этому — *принцип Паули*, согласно которому в каждом квантовом состоянии может находиться не более одного электрона.

Принцип Паули приводит к тому, что даже при абсолютном нуле температуры электроны не могут находиться в состоянии покоя, так как нельзя все электроны поместить в состояние с равным нулю импульсом. Оказывается, что уровню энергии с нулевым импульсом соответствуют только два различных квантовых состояния. Поэтому только два из N электронов могут иметь равные нулю импульс и энергию. Всем остальным электронам придется занимать состояния с отличным от нуля импульсом и энергией.

Импульс Ферми. При температуре $T = 0$ К система электронов будет иметь наименьшую возможную энергию, если все состояния с импульсами, меньшими некоторого предельного значения, называемого *импульсом Ферми* p_{F}, будут заняты, а все состояния с большими импульсами — пусты. Функция распределения электронов по состояниям $f(\mathbf{p})$, дающая вероятность заполнения данного состояния с импульсом $\mathbf{p}$, будет, следовательно, при $T = 0$ К равна единице при $|\mathbf{p}| \leqslant p_{\mathrm{F}}$ и равна нулю при $|\mathbf{p}| > p_{\mathrm{F}}$ (рис. 113).

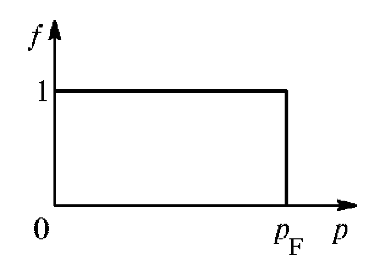


Рис. 113. Функция распределения электронов при абсолютном нуле температуры

Значение импульса Ферми зависит от объемной концентрации электронов в зоне проводимости $n = N/V$, где V — объем образца. Характер этой зависимости можно установить на основании соотношений неопределенностей Гейзенберга. Пусть образец металла имеет форму прямоугольного параллелепипеда со сторонами L_x, L_y, L_z. Для электрона, находящегося внутри образца, наибольшая неопределенность в значении x-координаты равна размеру образца в этом направлении; поэтому наименьшая неопределенность его x-компоненты импульса Δp_x определяется соотношением

$$\Delta p_x \sim \frac{h}{L_x}\,. \qquad (1)$$

Такие же соотношения справедливы и для Δp_y, Δp_z. Поэтому для произведения неопределенностей имеем

$$\Delta p_x \Delta p_y \Delta p_z \sim \frac{h^3}{L_x L_x L_x} = \frac{h^3}{V}. \tag{2}$$

Поскольку Δp_x есть наименьшая неопределенность значения p_x, совместимая с условием пребывания электрона внутри образца, то это — та наименьшая величина, на которую следует изменить значение p_x для того, чтобы можно было считать, что электрон находится в другом квантовом состоянии. Поэтому произведение $\Delta p_x \Delta p_y \Delta p_z$ дает объем ячейки в пространстве импульсов, в пределах которой при изменении импульса электрона его состояние не меняется. Это означает, что N электронов в соответствии с принципом Паули занимают в пространстве импульсов объем, пропорциональный $N\Delta p_x \Delta p_y \Delta p_z$. В соответствии с (2)

$$N\Delta p_x \Delta p_y \Delta p_z \approx N\frac{h^3}{V} = h^3 n. \tag{3}$$

В основном состоянии энергия системы электронов имеет наименьшее возможное значение. Энергия каждого электрона $\varepsilon(\mathbf{p}) = p^2/2m^*$, если эффективная масса m^* изотропна. Поэтому при $T = 0$ K будут заполнены все состояния в пространстве импульсов, лежащие внутри сферы радиуса p_F. Этот радиус можно найти, приравнивая объем части пространства импульсов, даваемой формулой (3), объему шара радиуса p_F с центром в начале координат: $(4/3)\pi p_F^3$. Поэтому $p_F^3 \approx h^3 n$, откуда

$$p_F \approx h n^{1/3}. \tag{4}$$

Энергия Ферми. Наибольшая энергия, которую может иметь электрон проводимости в металле при $T = 0$ K, это энергия электрона с $p = p_F$:

$$\varepsilon_F = \frac{p_F^2}{2m^*} = \frac{h^2 n^{2/3}}{m^*}. \tag{5}$$

Она называется *энергией Ферми*. Интересно сравнить энергию Ферми ε_F с характерной тепловой энергией kT. Для типичных металлов концентрация электронов проводимости составляет около 10^{21}–10^{23} см$^{-3}$, а эффективная масса m^* близка к массе свободного электрона. Поэтому по формуле (5) легко подсчитать, что ε_F оказывается порядка нескольких электронвольт. Легко убедиться, что тепловая энергия kT становится сравнимой с энергией Ферми только при температурах приближающихся к 50 000 K! Для всех металлов при всех температурах вплоть до точки плавления $kT \ll \varepsilon_F$.

Нетрудно подсчитать скорость электрона, обладающего энергией ε_F. Она оказывается равной примерно 10^6 м/с. Таким образом, вследствие принципа Паули электроны проводимости в металле при абсолютном нуле температуры отнюдь не покоятся. Но все же их максимальная скорость в сотни раз меньше скорости света, что оправдывает использование для электронов в металлах представлений нерелятивистской физики.

О вкладе электронов в теплоемкость металла. При повышении температуры от абсолютного нуля до некоторого значения T средняя энергия каждой частицы, согласно представлениям классической статистической механики, должна возрасти на величину порядка kT. Но для подчиняющихся принципу Паули электронов в металле это не так. При температуре абсолютного нуля электроны проводимости уже имеют энергию порядка ε_F, которая всегда много больше kT. Эта энергия не связана с тепловым движением электронов. При повышении температуры от 0 K до T отнюдь не каждый электрон может увеличить свою энергию на величину порядка kT, так как для большинства электронов все состояния, отстоящие от данного по энергии на величину порядка kT, заняты другими электронами.

Увеличить свою энергию смогут только электроны, занимающие состояния в интервале шириной примерно kT около ε_F. Поэтому тепловое движение при конечной температуре приводит к размытию ступеньки функции распределения f в узком интервале значений энергии порядка kT (рис. 114). Основная часть электронов будет занимать те же состояния, что и при абсолютном нуле. О таких свойствах функции распределения говорят как о *сильном вырождении* электронного газа в металлах.

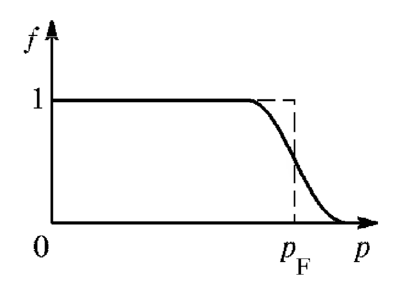


Рис. 114. Температурное размытие края функции распределения электронов в металле

Отличие функции распределения электронов проводимости в металле от максвелловской функции распределения классического газа объясняет особенности наблюдаемых электронных свойств металлов. Сравним, например, теплоемкости классического одноатомного газа и вырожденного газа электронов проводимости. В состоянии теплового равновесия в одноатомном классическом газе средняя энергия хаотического движения одной частицы равна $(3/2)kT$. Поэтому внутренняя энергия всего газа из N частиц равна $(3/2)NkT$, а его теплоемкость при постоянном объеме равна $(3/2)Nk$ и не зависит от температуры.

В вырожденном электронном газе в тепловом движении может участвовать лишь часть электронов порядка kT/ε_F, поскольку именно такая доля полного их числа N содержится в интервале шириной около kT вблизи энергии Ферми ε_F. Каждый из этих электронов приобретает дополнительную энергию порядка kT, и, следовательно, пол-

ная энергия хаотического теплового движения электронов E по порядку величины равна

$$E \approx N \frac{kT}{\varepsilon_F} kT. \tag{6}$$

По своему смыслу эта энергия представляет собой вклад электронов проводимости во внутреннюю энергию металла. Поэтому для теплоемкости электронного газа получаем

$$C_V = \frac{\Delta E}{\Delta T} \approx N \frac{k^2 T}{\varepsilon_F} \tag{7}$$

(с точностью до числового множителя порядка единицы).

Поскольку $kT/\varepsilon_F \ll 1$ (например, при комнатной температуре это отношение равно всего лишь 1/200), то теплоемкость вырожденного электронного газа примерно на два порядка меньше теплоемкости классического одноатомного газа с тем же числом частиц. Поэтому при комнатной температуре молярная теплоемкость металлов практически целиком определяется вкладом решетки и почти не отличается от молярной теплоемкости диэлектриков, в которых электроны проводимости отсутствуют.

Другое важное отличие теплоемкости электронного газа (7) заключается в том, что она пропорциональна температуре, в то время как теплоемкость классического газа от температуры не зависит.

Электропроводность металлов. Наряду со свойствами коллектива электронов проводимости, которые он проявляет в состоянии равновесия, большой интерес представляет изучение поведения электронов проводимости в металле в неравновесном состоянии, когда они движутся под действием приложенных внешних полей. Такие процессы называются *явлениями переноса*. Примерами таких явлений могут служить теплопроводность, электрическая проводимость, эффект Холла, термоэлектрические явления и т. д.

Остановимся подробнее на одном из наиболее важных неравновесных процессов — прохождении электрического тока в металле, описываемом законом Ома. Закон Ома является одним из самых ранних экспериментальных открытий в области электрического тока в веществе. Согласно этому закону плотность электрического тока пропорциональна напряженности приложенного электрического поля: $\mathbf{j} = \sigma \mathbf{E}$. Это уравнение описывает наблюдаемое поведение очень многих веществ в широком диапазоне условий.

Закон Ома нельзя вывести только из законов электродинамики, описывающих электрическое поле. Его можно получить на основе изучения процессов, происходящих в веществе при приложении электрического поля. Интересно отметить, что в разных веществах эти процессы могут сильно различаться, но, независимо от характера процессов рассеяния носителей связь между плотностью тока и

напряженностью поля при не слишком сильных полях всегда оказывается линейной.

В металлах сопротивление электрическому току обусловлено рассеянием электронов проводимости на тепловых колебаниях решетки и хаотически расположенных дефектах кристалла — в идеальной решетке сопротивление отсутствует. Феноменологически описать процессы рассеяния электронов проводимости можно путем введения так называемого *времени релаксации*.

Будем рассуждать следующим образом. В состоянии равновесия функция распределения электронов по импульсам сферически-симметрична. Поэтому средняя скорость электронов равна нулю и никакого направленного переноса заряда, т. е. электрического тока, нет.

Приложенное электрическое поле вызывает искажение функции распределения. В отличие от изотропного теплового воздействия, которое могло только размыть край распределения, не нарушая его сферической симметрии, направленное действие внешнего поля вызывает сдвиг всего распределения (рис. 115). На этом рисунке показано, как под действием электрического поля смещается в пространстве импульсов сфера радиуса p_F, внутри которой расположены занятые электронами состояния. Средняя заштрихованная область соответствует состояниям, которые заполнены электронами как до приложения электрического поля, так и после. Незаштрихованная часть соответствует состояниям, которые до приложения поля были заполнены, а после приложения поля оказались не занятыми. Наконец, дважды заштрихованная область соответствует состояниям, которые оказались заполненными после приложения поля. Видно, что теперь число электронов, движущихся в направлении действия силы электрического поля, больше, чем движущихся в противоположном направлении. Для такого сдвинутого в пространстве импульсов распределения средняя скорость уже отлична от нуля, что означает наличие тока в образце.

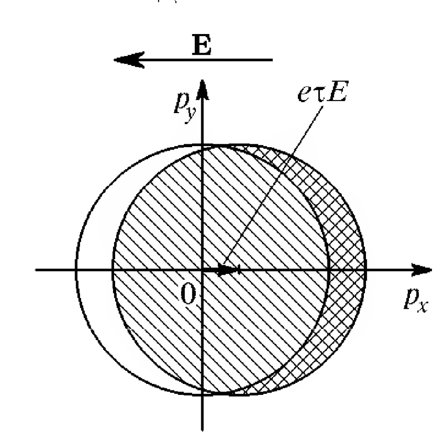


Рис. 115. Смещение сферической поверхности Ферми под действием приложенного электрического поля

Пусть в некоторый момент времени внешнее электрическое поле выключается. Очевидно, что система электронов будет возвращаться в состояние теплового равновесия. Предположим, что скорость приближения к равновесию в каждый момент пропорциональна величине отклонения от равновесия. Тогда и изменение средней скорости направленного движения электронов будет происходить по такому же закону. Так как в равновесии средняя скорость равна ну-

лю, то изменение скорости происходит пропорционально самой средней скорости:

$$\frac{d\langle v\rangle}{dt} = -\frac{\langle v\rangle}{\tau}. \tag{8}$$

Поскольку решение этого уравнения имеет вид

$$\langle v(t)\rangle = v_0 e^{-t/\tau}, \tag{9}$$

то τ есть время, в течение которого скорость направленного движения убывает в e раз. Оно называется *временем релаксации*.

Закон изменения средней скорости $\langle v(t)\rangle$ направленного движения электронов проводимости (9) имеет в точности такой же вид, что и уравнение движения тела в вязкой среде под действием силы трения, пропорциональной скорости:

$$\mathbf{F}_{\text{тр}} = -\frac{m}{\tau}\mathbf{v}.$$

Это наводит на мысль, что взаимодействие электронов проводимости с колебаниями решетки и примесями, приводящее к описываемой уравнением (9) релаксации средней скорости, можно учесть путем введения некоторой эффективной, непрерывно действующей на электрон силы трения $\mathbf{F}_{\text{тр}} = -\frac{m}{\tau}\mathbf{v}$.

При наличии постоянного электрического поля $\mathbf{E}$ на электрон, кроме силы трения, действует еще сила $-e\mathbf{E}$, где e — абсолютная величина заряда электрона. В стационарном состоянии электрон движется с такой постоянной скоростью, при которой эти силы уравновешиваются: $-e\mathbf{E} = \frac{m}{\tau}\mathbf{v}$. Такой же будет средняя скорость $\langle\mathbf{v}\rangle$ движения электронов в металле под действием поля $\mathbf{E}$: $\langle\mathbf{v}\rangle = -\frac{e\tau}{m}\mathbf{E}$.

Плотность тока $\mathbf{j}$, обусловленного направленным движением электронов со скоростью $\langle\mathbf{v}\rangle$, равна $-en\langle\mathbf{v}\rangle$. Подставляя сюда значение $\langle\mathbf{v}\rangle$, находим

$$\mathbf{j} = \frac{ne^2\tau}{m}\mathbf{E} = \sigma\mathbf{E}. \tag{10}$$

Это и есть закон Ома. Вычисление времени релаксации τ и, следовательно, удельной электрической проводимости σ требует детального рассмотрения механизмов рассеяния электронов проводимости.

Описанная выше физическая картина электропроводности металлов основывалась на предположении, что перенос заряда обусловлен только электронами зоны проводимости. Во многих металлах, однако, электропроводность определяется электронами нескольких энергетических зон. Например, электропроводность железа создается как электронами зоны проводимости, возникшей из $4s$-уровней изолированного атома железа, так и дырками из заполненной на 80 % энергетической зоны, возникшей из $3d$-уровней атома железа. Но и в случае дырочной проводимости остаются спра-

ведливыми развитые представления о роли рассеяния носителей тока для объяснения закона Ома.

Плазменные свойства металлов. При изучении электронных свойств металлов, таких, как теплоемкость и электропроводность, систему свободных электронов можно было рассматривать как совокупность невзаимодействующих между собой частиц. Все наблюдаемые свойства объяснялись взаимодействием электронов проводимости с колебаниями кристаллической решетки и дефектами ее структуры — примесями, вакансиями и т. д.

Однако существует целый ряд явлений, в которых взаимодействие электронов играет определяющую роль. В таких случаях говорят о проявлении электронами плазменных свойств.

Выше при изучении плазмы мы считали, что к частицам, из которых состоит плазма, можно применять законы классической физики. В газоразрядной и тем более в ионосферной плазме это действительно так, ибо там расстояние между электронами много больше их дебройлевской длины волны.

Но для плазмы, представляющей собой электроны в металлах или полупроводниках, ситуация совсем иная. Сильное вырождение электронного газа в металле означает, что среднее расстояние между электронами одного порядка с их дебройлевской длиной волны h/p_F для электронов с импульсом Ферми p_F. Строгая теория плазменных свойств металлов очень сложна, однако основные закономерности можно установить с помощью анализа размерностей.

Поведение электронов проводимости описывается квантовой механикой. Поэтому мы должны ожидать, что в формулах, выражающих свойства такой плазмы, наряду с параметрами e, m, n, kT появится постоянная Планка h. Теперь наряду с безразмерным параметром γ, даваемым формулой (7) § 33 и характеризующим классическую плазму, появится еще один безразмерный параметр Γ:

$$\Gamma = \frac{h^2 n^{2/3}}{mkT}. \tag{11}$$

Физический смысл этого параметра легко уяснить, обращаясь к формуле (5): Γ представляет собой отношение энергии Ферми ε_F к характерной тепловой энергии kT. Если $\Gamma \ll 1$, т. е. $\varepsilon_F \ll kT$, то электронный газ по своим свойствам близок к рассмотренной выше классической плазме. Так бывает в полупроводниках при не слишком больших концентрациях электронов проводимости и достаточно высоких температурах.

Как мы видели, для электронов в металлах при всех температурах вплоть до точки плавления $\varepsilon_F \gg kT$ и, следовательно, $\Gamma \gg 1$. Для такой плазмы потенциальную энергию кулоновского взаимодействия электронов $e^2 n^{1/3}$ следует сравнивать не с кинетической энергией теплового движения kT, а с энергией Ферми $\varepsilon_F = p_F^2/2m \approx h^2 n^{2/3}/m$. Сте-

пень близости электронной плазмы в металлах к идеальному газу характеризуется отношением энергии взаимодействия к энергии Ферми, т. е. безразмерным параметром, который принято обозначать через r_s:

$$r_s = \frac{e^2 n^{1/3}}{h^2 n^{2/3}/m} = \frac{e^2 m}{h^2 n^{1/3}}. \qquad (12)$$

Нетрудно видеть, что параметр r_s пропорционален отношению среднего расстояния между электронами $n^{-1/3}$ к боровскому радиусу $a_0 = \hbar/(me^2)$. Электронный газ в металле был бы близок к идеальному газу при $r_s \ll 1$.

Из формулы (12) также видно, что в противоположность классической плазме электронный газ тем ближе к идеальному газу, чем выше его концентрация. Конечно, при увеличении концентрации электронов n потенциальная энергия их взаимодействия возрастает. В случае классической плазмы, где кинетическая энергия не зависит от концентрации, плазма с ростом концентрации по своим свойствам все меньше напоминала бы идеальный газ.

Для электронов в металле, как мы видели, вследствие принципа Паули кинетическая энергия растет с увеличением концентрации как $n^{2/3}$. Поэтому, несмотря на увеличение потенциальной энергии, которая растет как $n^{1/3}$, отношение потенциальной энергии к кинетической убывает. Другими словами, относительная роль взаимодействия электронов при увеличении их концентрации становится все меньше и меньше. Оказывается, что для реальных металлов параметр r_s больше единицы, т. е. концентрация электронов все-таки недостаточно высока для того, чтобы электронный газ можно было считать идеальным. Для электронов в металле более подходящим является название «электронная жидкость», чем «электронный газ». Тем не менее последовательная теория электронной жидкости металлов показала, что многие свойства, полученные в модели свободных электронов, качественно остаются справедливыми и при учете электронного взаимодействия.

Как и для классической плазмы, для электронов проводимости в металле характерно существование плазменных колебаний, обусловленных кулоновским взаимодействием электронов и ионов кристаллической решетки. В выражение (3) § 33 для частоты плазменных колебаний ω_p не входит ни температура T, ни постоянная Планка $\hbar$. Это соответствует физической причине возможности существования плазменных колебаний: они обусловлены кулоновским взаимодействием заряженных частиц, а не тем, какими законами — классическими или квантовыми должно описываться движение.

Энергия кванта плазменных колебаний (плазмона) $\hbar\omega_p$ при характерных для металлов концентрациях электронов проводимости оказывается больше энергии Ферми ε_F — максимальной энергии

электронов для металла в основном состоянии. Поэтому плазменные колебания в металлах не могут возникать за счет энергии движения отдельных электронов, в противоположность классической плазме. Плазменные колебания в металлах могут возбуждаться только пролетающими через образец быстрыми заряженными частицами или падающими электромагнитными волнами. Роль плазменных колебаний в состоянии равновесия сводится по существу только к экранированию кулоновского взаимодействия электронов. Именно поэтому модель свободных электронов качественно верно объясняет многие электронные свойства металлов.

Частота плазменных колебаний ω_p представляет собой граничную частоту для электромагнитных волн, которые могут распространяться в плазме. Волны с частотой, меньшей ω_p, не могут проникать в плазму, так как низкочастотные электромагнитные поля таких волн экранируются заряженными частицами плазмы. Падающая на границу плазмы волна при $\omega < \omega_p$ отражается от границы. Если же частота электромагнитной волны выше плазменной частоты, то такая волна проникает в плазму.

Концентрация свободных электронов в щелочных металлах такова, что частота плазменных колебаний соответствует ультрафиолетовой области спектра. Поэтому в ультрафиолетовой области щелочные металлы прозрачны, хотя в видимой и инфракрасной областях спектра хорошо отражают падающее на них излучение. Напомним, что для ионосферной плазмы граница прозрачности попадает в диапазон метровых радиоволн.

- Поясните, почему распределение электронов проводимости в металле по разрешенным энергетическим уровням резко отличается от соответствующего распределения для газа классических частиц, хотя поведение отдельного электрона проводимости в кристалле сходно с поведением классической свободной частицы.
- Что такое импульс Ферми? Как эта величина связана с концентрацией электронов проводимости в металле?
- Что такое энергия Ферми? Как оценить скорость электрона в металле, обладающего энергией, равной энергии Ферми?
- Чем объясняется малый по сравнению с кристаллической решеткой вклад электронов проводимости в теплоемкость металлов?
- Поясните, почему теплоемкость вырожденного электронного газа пропорциональна температуре, в то время как теплоемкость классического газа от температуры не зависит.
- Что такое время релаксации для электронов в металле? Разъясните смысл этого понятия на примере затухания направленного движения электронов.
- Как объясняется закон Ома на основе представлений об эффективной силе трения, действующей на электроны проводимости при их направленном движении под действием электрического поля?

- Почему, в противоположность классической плазме, электронный газ в металлах с увеличением концентрации приближается по своим свойствам к идеальному газу?
- Покажите, что энергия кванта плазменных колебаний в реальном металле превосходит максимальную энергию отдельного электрона. К каким следствиям это приводит?
- Почему, строго говоря, свободные электроны в металле некорректно рассматривать как идеальный газ?
- Чем объясняется прозрачность щелочных металлов в ультрафиолетовой области спектра?

§ 36. Электронные свойства полупроводников

Как уже отмечалось, электрические свойства полупроводников очень сильно зависят от наличия и концентрации тех или иных примесей. Чтобы разобраться в этих вопросах, нужно иметь представление о природе и энергетическом спектре состояний электронов, связанных с атомами примесей.

Роль примесей в полупроводниках. Наибольшее влияние на свойства полупроводников оказывают так называемые *донорные* и *акцепторные примеси* (или *доноры* и *акцепторы*). О донорах мы уже упоминали в § 34 при классификации полупроводников. Наличие донорных примесей в кристалле приводит к появлению дополнительных разрешенных электронных уровней энергии, лежащих в запрещенной зоне вблизи дна зоны проводимости. Акцепторные примеси приводят к появлению дополнительных уровней в запрещенной зоне, лежащих вблизи потолка валентной зоны. Важно, что в обоих случаях примесные уровни отделены от соответствующих разрешенных зон энергетическими интервалами, малыми по сравнению с шириной запрещенной зоны.

Доноры. В качестве простого примера донорного центра в полупроводнике рассмотрим кристалл кремния, в котором один из узлов решетки занят атомом фосфора (рис. 116). Кремний кристаллизуется в решетке типа алмаза, где каждый атом имеет четыре ближайших соседа, расположенных в вершинах правильного тетраэдра, в центре которого находится рассматриваемый атом. Четыре электрона из внешней электронной оболочки каждого атома участвуют в образовании направленных ковалентных связей с четырьмя окружающими его такими же атомами кремния. В картине энергетических зон кристалла кремния эти электроны занимают квантовые состояния валентной зоны, причем при абсолютном нуле температуры эта зона оказывается полностью заполненной.

Примесной атом фосфора, занимающий соседнюю ячейку в таблице Менделеева, имеет пять электронов во внешней оболочке. Четыре из них участвуют в образовании ковалентных связей с соседними атомами кремния. Пятый электрон фосфора оказывается лишним, он не нужен для образования ковалентных связей. Этот электрон не вполне свободен, потому что ядро атома фосфора обладает положительным зарядом, на единицу большим, чем ядро атома кремния. В результате этот электрон движется в кристалле по большой орбите вокруг примесного атома, подобно тому, как движется электрон вокруг ядра в атоме водорода. На рис. 116 эта орбита показана. Размеры такой орбиты велики по сравнению с межатомным расстоянием в кристалле.

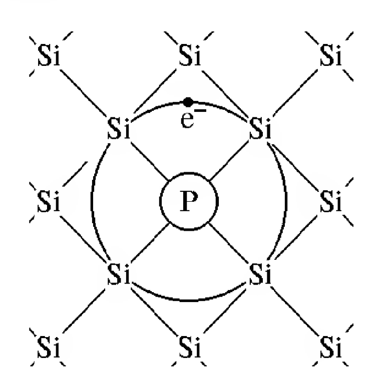


Рис. 116. Схематическое двумерное изображение донорного центра в кристалле кремния

Большой размер орбиты и малая энергия связи такого электрона объясняются тем, что сила взаимодействия его с ядром атома фосфора ослаблена диэлектрическими свойствами среды кристаллической решетки кремния, и тем, что эффективная масса электрона в зоне проводимости кремния меньше массы свободного электрона.

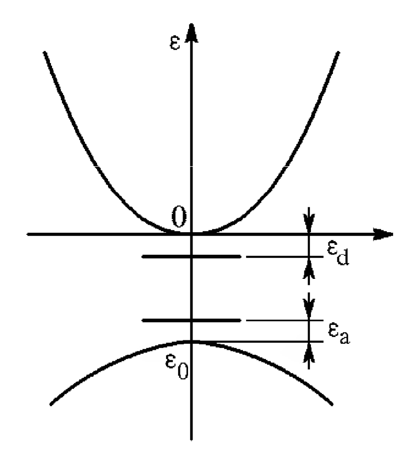


Рис. 117. Уровни энергии доноров и акцепторов

Таким образом, электрически нейтральное состояние примесного центра можно представить как локализованное состояние электрона (электрон связан с примесным центром), причем энергия этого состояния на диаграмме электронных уровней энергии кристалла лежит в запрещенной зоне вблизи дна зоны проводимости (рис. 117). Из такого локализованного состояния электрон может быть переведен в нелокализованное состояние зоны проводимости, если ему сообщить дополнительную энергию, не меньшую ε_d. При этом донорный центр становится ионизованным, так как на нем локализован избыточный положительный заряд.

Энергия электрона на донорном центре составляет около 0,01 эВ, поэтому при обычных температурах практически все донорные центры ионизованы.

Акцепторы. В кристаллах кремния и германия донорами могут служить не только атомы фосфора, но и атомы других элементов пятой группы таблицы Менделеева. В противоположность этому примеси элементов третьей группы в таких кристаллах являются акцепторами. В этом случае при замещении одного из основных атомов кристалла, например атомом бора, для образования направ-

ленных ковалентных связей с соседями не хватает одного электрона.

Представляется естественным, что одна из четырех валентных связей, направленных к бору, должна остаться ненасыщенной. Это соответствует отсутствию электрона в заполненной валентной зоне, т. е. положительной дырке. Пространственное положение отсутствующей связи может изменяться, переходя от одного из соседних с примесью атомов кремния к другому.

Наиболее просто такую ситуацию можно описать, используя представление о связанной дырке, движущейся в кулоновском поле, создаваемом эффективным отрицательным зарядом $-e$, происхождение которого связано с тем, что заряд ядра примесного атома на единицу меньше заряда ядра атома кремния. Как и в случае доноров, это поле ослаблено диэлектрической проницаемостью кристаллической решетки. Поскольку и для дырки эффективная масса в данном случае меньше массы свободного электрона, то радиус боровской орбиты дырки, движущейся вокруг примесного центра, простирается на большое (по сравнению с постоянной решетки) расстояние. Энергия связи дырки невелика, поэтому соответствующий акцепторному центру энергетический уровень электрона лежит в запрещенной зоне немного выше потолка валентной зоны (рис. 117).

Дырка, связанная с акцепторным центром, может быть переведена из локализованного состояния в нелокализованное состояние валентной зоны, если ей сообщить дополнительную энергию, превосходящую энергию ионизации акцепторного центра ε_a. Акцепторный центр при этом становится отрицательно заряженным.

Энергия связи дырки на акцепторном центре, как и в случае донорных центров, также измеряется сотыми долями электронвольта. Поэтому при обычных температурах и акцепторные центры практически полностью ионизованы.

Электропроводность полупроводников. Проводимость кристалла, обусловленная наличием доноров, поставляющих электроны в зону проводимости, называется проводимостью *n*-типа (от англ. negative — отрицательный), а проводимость, обусловленная акцепторными примесями, поставляющими положительные дырки в валентную зону, — проводимостью *p*-типа (от англ. positive).

Электропроводность металла при изменении внешних условий меняется только из-за изменения времени релаксации, которое зависит от температуры и концентрации дефектов кристаллической решетки. Однако число носителей зарядов — свободных электронов — остается постоянным. В полупроводнике дело обстоит совершенно иначе. Главная причина изменения электропроводности полупроводника заключается в изменении концентрации носителей заряда.

Собственный полупроводник при абсолютном нуле температуры, как и диэлектрик, не содержит свободных носителей, так как вален-

тная зона заполнена целиком, а зона проводимости пуста. С повышением температуры часть электронов переходит из валентной зоны в зону проводимости. Концентрация свободных носителей при этом зависит от отношения характерной тепловой энергии kT к ширине запрещенной зоны.

В примесных полупроводниках появление свободных носителей при низких температурах связано в основном не с переходом электронов в зону проводимости из валентной зоны, а с ионизацией локализованных электронных состояний, энергии которых лежат в запрещенной зоне.

Реальные кристаллы всегда содержат примеси. Поэтому собственная проводимость преобладает только в очень чистых полупроводниковых кристаллах при сравнительно высоких температурах. При глубоком охлаждении собственный полупроводник неизбежно должен стать примесным.

Распределение носителей тока по энергиям. Как распределены электроны в зоне проводимости полупроводника и дырки в валентной зоне по уровням энергии? Это нужно знать для построения микроскопической теории электропроводности и других наблюдаемых свойств, обусловленных электронами. Качественно ответ на этот вопрос можно дать, основываясь на значении безразмерного параметра Г, определяемого формулой (11) предыдущего параграфа:

$$\Gamma = \frac{h^2 n^{2/3}}{m^* kT}. \qquad (1)$$

Легко видеть, что этот параметр представляет собой квадрат отношения дебройлевской длины волны частицы h/p, имеющей характерный тепловой импульс $p = \sqrt{m^* kT}$, к среднему расстоянию между частицами $\langle r \rangle \approx n^{-1/3}$. При $\Gamma \ll 1$ поведение носителей тока определяется законами классической физики и их функция распределения по энергиям и скоростям такая же, как у газа классических частиц. При $\Gamma \gg 1$ необходимо учитывать принцип Паули, ибо среднее расстояние между частицами много меньше дебройлевской длины волны. В этом случае, как и у металлов, распределение носителей тока по уровням энергии имеет вид ступеньки: почти все состояния с импульсами, меньшими импульса Ферми, заполнены; с большими импульсами — свободны. Все физические свойства при этом обусловлены вкладом носителей с энергиями, соответствующими области размытия этой ступеньки.

В полупроводниках со смешанной проводимостью, где одновременно есть и электроны в зоне проводимости, и дырки в валентной зоне, параметр Г для каждого типа носителей имеет свое значение. Может быть, например, так, что распределение электронов описывается ступенчатой функцией, а дырок — максвелловской. В обыч-

ных условиях в полупроводниках, используемых во многих электронных устройствах, распределение носителей тока описывается классической максвелловской функцией.

Эффект Холла. Представление о дырках как о свободных носителях тока с положительными зарядами было введено для того, чтобы наглядно описать вклад в электропроводность валентной зоны, почти целиком заполненной электронами. Убедительное доказательство реальности положительных дырок дает *эффект Холла*, открытый в 1879 г., задолго до возникновения представлений о самих дырках.

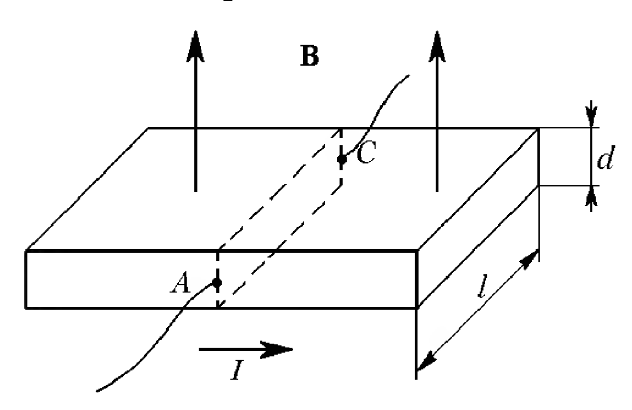


Рис. 118. К объяснению эффекта Холла

Эффект Холла можно наблюдать при прохождении электрического тока в любых твердых проводящих телах, помещенных во внешнее магнитное поле. Он заключается в появлении внутри проводящего тела электрического поля, напряженность **E** которого перпендикулярна как индукции **B** внешнего магнитного поля, так и плотности **j** пропускаемого через образец тока.

Рассмотрим прямоугольную пластинку из проводящего вещества, помещенную в однородное магнитное поле с индукцией **B** (рис. 118). Пусть ток I пропускается через образец слева направо, как показано на рисунке, и осуществляется одним типом носителей с зарядом q, который может быть как положительным, так и отрицательным. Обозначив концентрацию носителей через n, а их дрейфовую скорость — через v, запишем выражение для силы тока I в виде

$$I = jS = qnvld. \tag{2}$$

При наложении магнитного поля **B** на движущиеся в пластине носители тока действует сила Лоренца

$$F_{\text{Л}} = \frac{q}{c} vB, \tag{3}$$

направленная перпендикулярно как магнитному полю **B**, так и скорости **v** направленного движения носителей. В результате на боковых гранях пластины скапливаются заряды противоположного знака, создающие электрическое поле **E**. Действующая со стороны этого поля сила

$$F = qE \tag{4}$$

уравновешивает силу Лоренца, обеспечивая направленное движение носителей тока параллельно боковым граням пластины. Наличие

электрического поля **E**, перпендикулярного направлению движения зарядов, означает, что между точками A и C на противоположных гранях образца, которые в отсутствие магнитного поля имели одинаковый потенциал, теперь появляется некоторая разность потенциалов U, пропорциональная напряженности поля E:

$$U = Ed. \tag{5}$$

Приравнивая правые части выражений (3) и (4) и подставляя в получившееся равенство v из (2) и E из (5), находим

$$U = \frac{1}{nqc}\frac{I}{d}. \tag{6}$$

Формула (6) справедлива как для положительных, так и для отрицательных носителей заряда. Зависящий только от свойств материала пластины (но не от ее размеров) коэффициент пропорциональности в (6)

$$R = \frac{1}{nqc} \tag{7}$$

называется *постоянной Холла*. Знак постоянной Холла совпадает со знаком носителей заряда. Поэтому для образцов с электронной и дырочной проводимостью постоянная Холла имеет противоположные знаки.

Измерения постоянной Холла недвусмысленно подтвердили существование полупроводников как с электронной, так и с дырочной проводимостью. Более того, коэффициент Холла оказался положительным и для некоторых металлов, например для железа, что обусловлено 80%-ным заполнением энергетической зоны, возникшей из $3d$-уровня энергии изолированного атома железа.

- Почему электрические свойства полупроводников очень сильно зависят от наличия примесей?
- Поясните аналогию между локализованным состоянием электрона на примесном центре и состоянием электрона в атоме водорода. Почему радиус орбиты электрона, связанного с примесным центром, много больше боровского радиуса $a_0 = \hbar^2/(me^2)$?
- Что является аналогом ионизации атома водорода для нейтрального донорного примесного центра в кристалле?
- Опишите физическую модель примесного акцепторного центра в полупроводниковом кристалле. Почему в рамках такой модели энергетические уровни локализованных электронных состояний лежат в запрещенной зоне вблизи потолка валентной зоны?
- Как объяснить ионизацию акцепторного центра, не прибегая к концепции положительных дырок, а рассматривая переход электрона из состояния в валентной зоне на примесный уровень?
- В чем состоит принципиальное различие причин изменения электропроводности металлов и полупроводников при изменении внешних условий?

- Поясните, почему в реальных образцах собственных полупроводников при очень низких температурах преобладает примесная проводимость.
- Почему при низкой концентрации носителей тока в полупроводниках их распределение по энергиям будет классическим, а при высокой концентрации функция распределения имеет вид ступеньки? Как изменится функция распределения при повышении температуры?
- Что такое эффект Холла? Каким образом он позволяет судить о знаке носителей заряда в исследуемом образце?
- Объясните, почему на противоположных боковых гранях образца в магнитном поле при пропускании тока скапливаются заряды противоположного знака, хотя электрический ток переносится зарядами одного знака.

§ 37. Полупроводниковые приборы

Полупроводниковые материалы широко применяются в современной электронике. Полупроводниковые приборы — диоды, транзисторы, интегральные схемы — выполняют те же функции, что и электронные вакуумные лампы. Но их габариты и потребляемая энергия настолько меньше, что они сделали возможным решение принципиально новых задач. Например, современный компьютер содержит столько функциональных элементов, что само его существование возможно только благодаря очень малому размеру этих элементов: на кремниевом монокристалле размером около 6,5 мм умещается несколько десятков тысяч транзисторов, соединенных в сложную электрическую схему.

Почти все полупроводниковые приборы имеют в основе монокристалл, содержащий две или более области с различными примесями. На границах областей с разным легированием возникают так называемые *переходы*. Перепады концентрации примесей в переходе и возникающие благодаря этому электрические поля вызывают целый ряд существенно нелинейных эффектов в электропроводности, благодаря которым и функционируют полупроводниковые приборы.

***p–n*-переходы.** Принцип действия большинства приборов может быть понят на основе простейшего перехода, называемого *p—n-переходом*. Этот переход представляет собой систему из двух полупроводников, находящихся в контакте, причем в одном из них в качестве носителей преобладают электроны (*n*-тип), а в другом — дырки (*p*-тип).

Остановимся подробнее на качественном описании свойств *p—n*-перехода. Такой переход может быть изготовлен множеством способов, например диффузией донорных примесей с одного конца и акцепторных примесей с противоположного конца германия с собственной проводимостью, так чтобы возник переход от *n*- к *p*-области

в очень тонком слое порядка 10^{-4} см. Схема такого *p*—*n*-перехода показана на рис. 119.

Если бы *p*- и *n*-области не были в контакте друг с другом, то каждая из них содержала бы определенное число неподвижных доноров и акцепторов, показанных на рис. 119*а* кружками с плюсами и минусами в них, а также носители — электроны и дырки, показанные знаками минус и плюс (без кружочков). Так как эти не соприкасающиеся области электрически нейтральны, то в каждой из них число носителей заряда равно числу атомов примеси. Конечно, благодаря тепловому движению в каждой из областей могут возникать и электронно-дырочные пары (собственная проводимость), но на этом рисунке они для простоты не показаны.

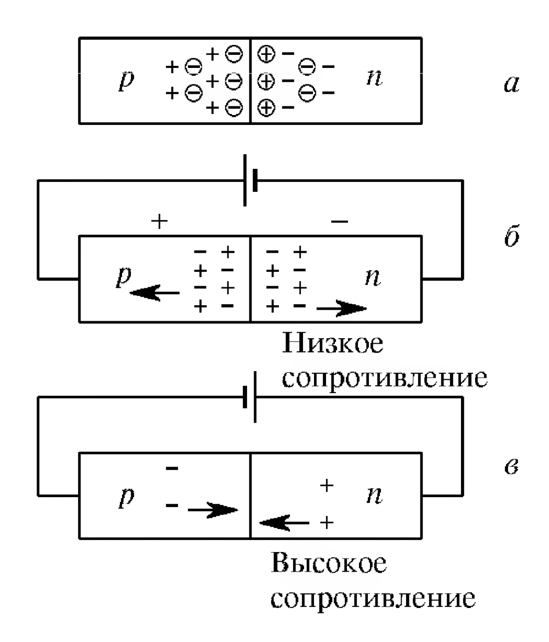


Рис. 119. Схема *p*—*n*-перехода, в котором *p*- и *n*-области монокристалла отличаются только типом примесей

Диффузия электронов и дырок. Основные и неосновные носители. Поскольку на самом деле *p*- и *n*-области перехода находятся в контакте, то действительная картина отличается от нарисованной выше. Наиболее важное отличие определяется диффузией электронов из правой *n*-области в левую *p*-область и диффузией дырок во встречном направлении. В результате этой диффузии нарушается электронейтральность у границ *p*- и *n*-областей и в переходе возникает электрическое поле, препятствующее дальнейшей диффузии электронов и дырок через переход. Устанавливается динамическое равновесие, при котором в прилегающих к переходу областях наряду с *основными носителями* (электронами в *n*-области и дырками в *p*-области) появляются *неосновные носители* с зарядами противоположного знака (дырки в *n*-области и электроны в *p*-области). Возникшее в переходе электрическое поле представляет собой потенциальный барьер для основных носителей и, напротив, «крутой склон» для неосновных, стремящийся вернуть их в ту область, где они были основными.

Вольт-амперная характеристика. Это динамическое равновесие нарушается при включении *p*—*n*-перехода во внешнюю электрическую цепь, причем результаты оказываются совершенно различными в зависимости от полярности приложенного к переходу напряжения (рис. 119*б*, *в*). Если полярность напряжения такая, как показано на рис. 119*б*, то электрический ток через *p*—*n*-переход

обусловлен движением основных носителей каждой области: вклад в ток дают электроны, пересекающие переход справа налево, и дырки, пересекающие его слева направо. При такой полярности приложенного напряжения потенциальный барьер в переходе понижается. Чем больше приложенное напряжение, тем значительнее понижение барьера, и электронам и дыркам легче диффундировать через переход. Поэтому, в отличие от однородного образца, где справедлив закон Ома, здесь зависимость тока от напряжения имеет ярко выраженный нелинейный характер.

Если приложенное напряжение имеет противоположную полярность, как показано на рис. 119*в*, электрический ток через *p*—*n*-переход обусловлен движением неосновных носителей каждой области: вклад в этот ток дают электроны, пересекающие переход слева направо, и дырки, пересекающие его справа налево. Теперь потенциальный барьер в переходе увеличивается с ростом приложенного напряжения. Однако для неосновных носителей это не барьер, а крутой спуск. Поэтому при такой полярности ток через переход, хотя и существенно меньше тока при «прямом» приложенном напряжении, но все же он растет при увеличении такого «обратного» напряжения. Схематически вид вольт-амперной характеристики *p*—*n*-перехода показан на рис. 120.

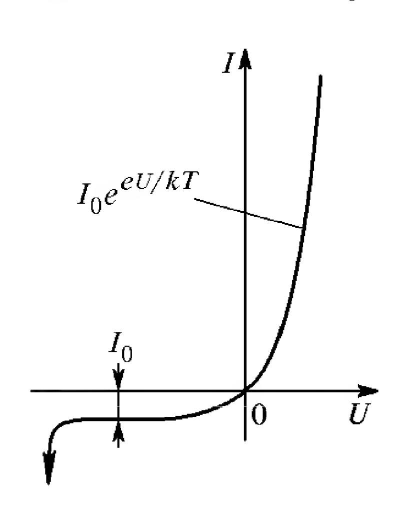


Рис. 120. Вольт-амперная характеристика *p*—*n*-перехода

Экспоненциальный характер роста силы тока при прямом напряжении связан с тем, что преодоление основными носителями потенциального барьера обусловлено их тепловым движением: барьер могут преодолеть только достаточно «быстрые» носители. Число носителей с энергией, превышающей высоту барьера, при максвелловском распределении экспоненциально зависит от отношения высоты барьера к характерной тепловой энергии kT. Такая экспоненциальная вольт-амперная характеристика типична для идеального *p*—*n*-перехода с плоской границей между *p*- и *n*-областями.

При увеличении обратного приложенного напряжения рост силы тока с напряжением быстро замедляется и он достигает насыщения, причем значение силы тока насыщения относительно невелико (по сравнению с силой тока при прямом напряжении). Малое значение тока насыщения обусловлено тем, что число неосновных носителей, осуществляющих перенос заряда при обратном напряжении, невелико и уже при сравнительно небольшом напряжении все они пересекают область *p*—*n*-перехода. При больших значениях обратного напряжения наступает пробой, т. е. лавинное нарастание тока, приводящее к разрушению *p*—*n*-перехода.

Нелинейный, асимметричный характер зависимости силы тока через *p*—*n*-переход от приложенного напряжения позволяет исполь-

зовать такой полупроводниковый диод для выпрямления переменного тока, детектирования (демодуляции) и других преобразований электрических сигналов.

Транзистор. Добавляя в монокристалл еще один *p*—*n*-переход или несколько таких переходов, можно получить *транзисторы*, четырехслойные диоды и другие приборы, позволяющие с помощью слабых сигналов управлять протеканием электрического тока.

Транзистор — это трехэлектродный прибор, аналогичный вакуумной электронной лампе — триоду. Он получается при последовательном включении двух *p*—*n*-переходов, например в конфигурации *n*—*p*—*n* (рис. 121). Узкий слой с проводимостью *p*-типа как бы за-

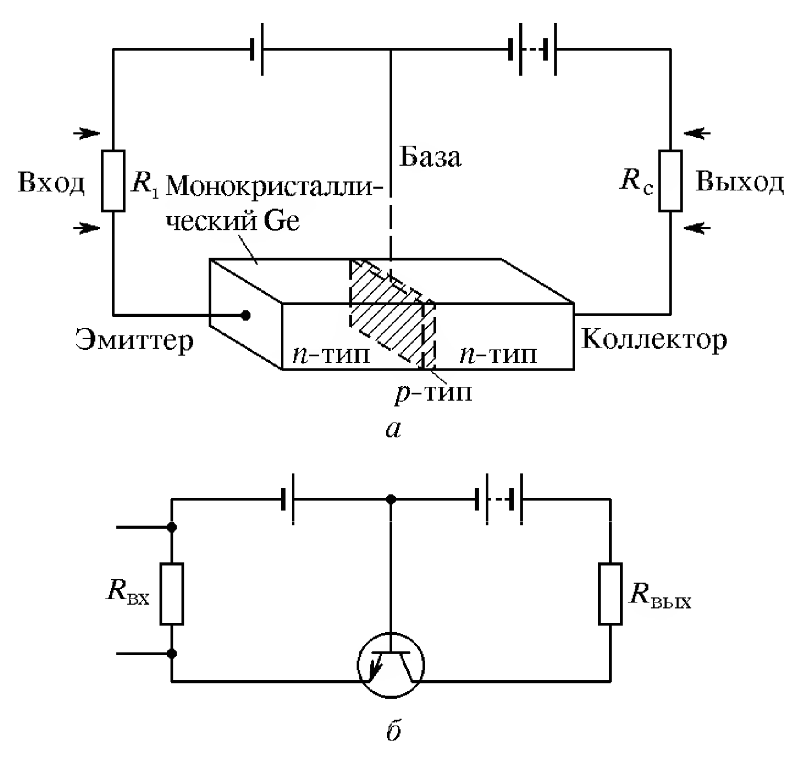


Рис. 121. Схематическое изображение транзистора, в котором две *n*-области монокристалла разделены узкой областью с проводимостью *p*-типа (*а*) и условное изображение такого транзистора на электрических схемах (*б*)

жат между двумя слоями с электронной проводимостью. Эту тонкую область между двумя *p*—*n*-переходами называют *базой* транзистора, а примыкающие к ней области *эмиттером* и *коллектором*. При включении транзистора в электрическую цепь по схеме, показанной на рис. 121, эмиттером служит левый *n*—*p*-переход, а коллектором — правый *p*—*n*-переход.

Транзисторное усиление слабого сигнала возникает за счет того, что слабый сигнал в цепи эмиттер-база вызывает сильное изменение большого тока, протекающего между эмиттером и коллектором.

Усилитель на транзисторе. Транзистор может быть включен в электрическую цепь многими способами. Простейшая схема с общей базой наиболее удобна для анализа работы транзистора. К эмиттеру

(левому переходу на рис. 121) приложено прямое постоянное напряжение около 1,5 В. Сопротивление $R_{вх}$ порядка 50 Ом. На него и подается переменный сигнал, который требуется усилить. К коллектору приложено более высокое обратное напряжение, например 9 В. Сопротивление $R_{вых}$ в выходной цепи имеет порядок 10^4 Ом. С него снимается усиленный сигнал.

Электроны, продиффундировавшие из эмиттера в базу, становятся там неосновными носителями. В области слабого поля между *n—p*- и *p—n*-переходами они движутся хаотически и могут оказаться вблизи *p—n*-перехода, отделяющего базу от коллектора. Здесь они оказываются в области сильного поля, затягивающего их в коллекторный переход. При этом электроны приобретают дополнительную энергию от поля, так как спускаются по «крутому склону» этого перехода. Проходя через коллекторную цепь, они создают напряжение на выходном сопротивлении $R_{вых}$, которое представляет собой усиленный отклик на сигнал, приложенный к базе.

При достаточно тонкой базе большая доля электронов, проникших в базу через первый *n—p*-переход (99 % и более), проходит через коллектор. Поэтому эффективность прибора велика. Коэффициент усиления можно оценить следующим образом. Почти все напряжение в цепи эмиттера приходится на входное сопротивление $R_{вх}$, поэтому ток через эмиттер можно записать как $U_{вх}/R_{вх}$. Если весь этот ток проходит через коллектор, то напряжение на выходном сопротивлении будет превышать напряжение на входном в $R_{вых}/R_{вх}$ раз. Во столько же раз возрастает мощность выходного сигнала. Источником мощности, выделяющейся на выходном сопротивлении, является батарея, включенная в цепь коллектора.

Транзистор с *p—n—p*-структурой работает так же, как только что описанный *n—p—n*-транзистор, но только здесь перенос заряда из цепи эмиттера в цепь коллектора через область базы осуществляется положительными дырками. Поэтому все напряжения здесь имеют обратную полярность.

Интегральные схемы. В различных электронных приборах транзисторы, диоды, резисторы, конденсаторы соединяются в сложные электрические цепи. Развитие технологии в течение последних десятилетий позволило создать так называемые большие интегральные схемы, в которых транзисторы, диоды, сопротивления и емкости можно получить в одном монокристалле (пластинке кремния или германия), формируя микроскопически малые *p*- и *n*-области. Затем эти элементы связывают друг с другом, металлизируя соответствующие участки поверхности кристалла. Конденсаторы в таком кристалле делаются с помощью оксидных покрытий, наращиваемых на поверх-

ности. Интегральная схема на одном монокристалле может выполнять функции целого прибора, например процессора в компьютере.

Светодиоды. Полупроводниковые приборы, кроме электрических цепей, находят еще и целый ряд других применений. Важнейшими из них являются оптические применения: светоизлучающие диоды и лазеры на *p—n*-переходах, солнечные батареи, детекторы излучения и генераторы микроволнового излучения.

Оказывается, что *p—n*-переходы, изготовленные на основе соединений элементов третьей и пятой групп таблицы Менделеева A^3B^5, таких, как арсенид галлия GaAs, фосфид галлия GaP, антимонид индия InSb, могут преобразовывать низковольтный электрический ток в свет с довольно высокой эффективностью. Это явление, называемое электролюминесценцией, было открыто еще в начале XX века в кристаллах карбида кремния SiC.

Наиболее эффективная электролюминесценция может быть осуществлена в *p—n*-переходе, где, пропуская ток через контакт, можно создать высокую концентрацию неосновных носителей. Оптическое излучение возникает при рекомбинации электрона и дырки, т. е. в процессе перехода электрона из зоны проводимости в одно из свободных состояний валентной зоны. Энергия испускае-

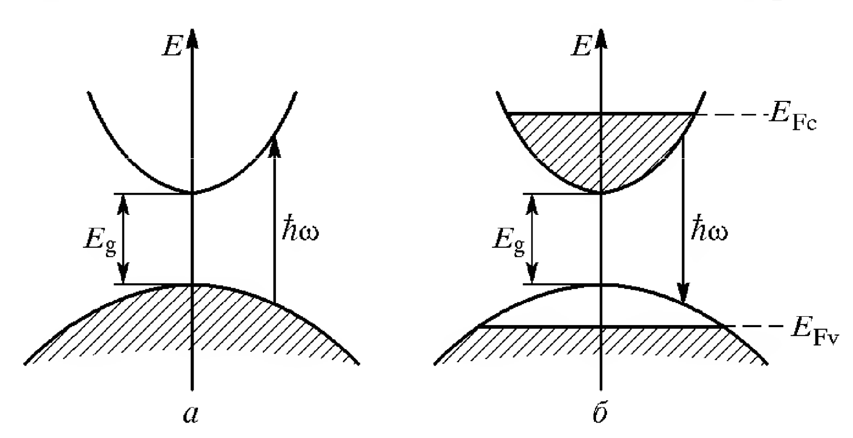


Рис. 122. Схема энергетических зон в кристаллах A^3B^5. На рис. *а* показано равновесное распределение электронов по уровням энергии, при котором свет поглощается, а на рис. *б* — сильно неравновесное распределение электронов с инверсией заселенностей, необходимое для усиления света

мого при этом светового кванта и, следовательно, частота излучения определяются разностью энергий электрона в начальном и в конечном состояниях. Эта энергия несколько больше ширины запрещенной зоны, а соответствующая частота у большинства используемых кристаллов приходится на видимую или ближнюю инфракрасную область спектра. Так, светодиоды на основе твердого раствора $GaAs_{1-x}P_x$ испускают красный свет, а на основе GaP с примесью азота — зеленый свет.

Выбор материалов типа соединений A^3B^5 продиктован тем, что у них дно зоны проводимости и потолок валентной зоны расположены в одной точке пространства импульсов при $\mathbf{p} = 0$ (рис. 122). Этим

они отличаются от германия и кремния, у которых минимумы зоны проводимости расположены при отличных от нуля значениях квазиимпульса, в то время как максимум валентной зоны приходится на значение $\mathbf{p} = 0$.

В элементарном акте поглощения фотона с образованием пары электрон—дырка и в обратном процессе рекомбинации электрона и дырки с испусканием фотона наряду с законом сохранения энергии должен выполняться закон сохранения импульса. Импульс фотона $\hbar\omega/c$ настолько мал по сравнению с характерными значениями квазиимпульсов электрона и дырки при данной температуре, что на рис. 122 оптические переходы изображаются вертикальными стрелками. Например, при рекомбинации электрон может перейти из зоны проводимости только в такое не занятое состояние валентной зоны, которому соответствует такое же значение квазиимпульса, какое было у него в зоне проводимости.

Полупроводниковые лазеры. Спонтанные переходы электронов из зоны проводимости на свободные места в валентной зоне в светодиодах сопровождаются некогерентным оптическим излучением, так же, как и спонтанные переходы изолированного атома из возбужденного состояния в основное. Для получения когерентного оптического излучения, т. е. для создания лазера, нужно обеспечить условия преобладания вынужденного излучения над спонтанным.

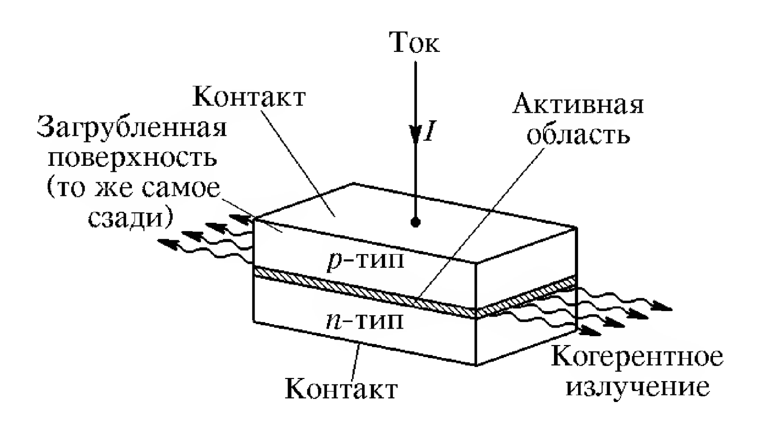


Рис. 123. Схема устройства полупроводникового лазера

В полупроводниковом лазере, так же, как и в газовом, когерентное вынужденное излучение возникает при инверсии заселенностей энергетических уровней валентной зоны и зоны проводимости, как показано на рис. 122*б*. Такую инверсию заселенностей получают путем накачки электронов пропусканием тока через p—n-переход. Обратная связь излучения с активной средой, необходимая для возникновения лазерной генерации, реализуется благодаря тщательно отполированным параллельным боковым граням монокристалла, образующим оптический резонатор (рис. 123).

Фотодиоды. Полупроводниковые приборы, в которых используется обратный процесс преобразования энергии излучения в электрическую энергию, называются *фотодиодами*. Они находят широкое применение в солнечных батареях.

Роль процессов рекомбинации. Приведенное выше описание принципа действия полупроводниковых приборов содержит ряд упрощений. Главное упрощение заключается в том, что не рассматривалась роль процессов рекомбинации электронов и дырок. Между тем именно рекомбинация определяет многие характерные черты электронных явлений в полупроводниках.

В некоторых полупроводниковых цепях, например в $p—n$-переходе, при прохождении электрического тока происходит смена типа носителей тока вдоль цепи. Если в n-области ток переносят главным образом электроны проводимости, то в p-области этот же ток переносят главным образом дырки валентной зоны. Электроны и дырки движутся в такой цепи навстречу друг другу, приближаются к $p—n$-переходу и, пересекая его, становятся неосновными носителями. Повышенная концентрация неосновных носителей в примыкающих к $p—n$-переходу участках приводит к интенсивной рекомбинации электронов и дырок.

Аналогичный рекомбинационный механизм «передачи эстафеты» от носителей одного типа к другим имеет место вблизи контакта подводящего металлического проводника с полупроводником p-типа.

Явное рассмотрение роли процессов рекомбинации позволяет объяснить основные особенности действия транзисторов. Для нормальной работы транзистора необходимо выполнение двух важных условий. Прежде всего база должна быть достаточно тонкой. Кроме того, концентрация примесей в эмиттере должна намного превосходить их концентрацию в базе.

О толщине и легировании базы транзистора. Тонкая база желательна по нескольким причинам. Во-первых, для эффективной работы необходимо, чтобы в коллектор попали все те носители, которые покинули эмиттер. Но некоторые из носителей неизбежно будут диффундировать к электрическому контакту базы вместо того, чтобы, пересекая базу, доходить до коллектора. Доля электронов, нашедших дорогу к базовому контакту, зависит от геометрии базы и тем меньше, чем тоньше база. Этот ток, который на языке вакуумных ламп можно назвать сеточным током и который должен быть как можно меньше по сравнению с током в цепи эмиттера (катода), тем меньше, чем тоньше база.

Во-вторых, с уменьшением толщины базы падает вероятность того, что эмитированный электрон рекомбинирует в базе, прежде чем продиффундирует к коллектору. Поскольку ток транзистора в базе создается неосновными носителями, их рекомбинация весьма вероятна и должна быть сведена к минимуму.

В-третьих, характеристики транзистора ухудшаются при высоких частотах сигналов и предельная частота определяется в основном толщиной базы. Прибор может удовлетворительно работать только до тех пор, пока время диффузии носителей через базу меньше периода колебаний усиливаемого сигнала: если напряжение поменяет знак прежде, чем большинство эмитированных электронов пересечет базу, они уже не дадут вклада в ток цепи коллектора.

Легко понять и причину того, почему степень легирования базы должна быть существенно ниже, чем у эмиттера. Во-первых, чем выше степень легирования базы n—p—n-транзистора, тем больше она содержит дырок. Следовательно, и вероятность того, что электрон рекомбинирует с дыркой прежде, чем достигнет коллектора, больше, а это нежелательно. Во-вторых, когда к эмиттеру приложено напряжение в прямом направлении, возникает не только поток электронов из эмиттера в базу, но и дырочный поток из базы в эмиттер. Чтобы паразитный поток из базы в эмиттер был много меньше потока из эмиттера, степень легирования базы должна быть малой.

Рекомбинация в лазере. Рекомбинация электронов и дырок играет большую роль и в работе полупроводникового лазера, где инверсная заселенность достигается пропусканием тока через p—n-переход, приводящего к увеличению концентрации неосновных носителей по обе стороны перехода. Вблизи p—n-перехода идет рекомбинация электронов и дырок. При низких значениях силы тока преобладающей может быть безызлучательная рекомбинация, но когда неравновесные концентрации неосновных носителей становятся достаточно большими, все большее значение приобретает рекомбинация с испусканием фотонов, приводящая к некогерентному оптическому излучению. При очень больших концентрациях неосновных носителей, когда создается инверсия заселенностей и вынужденное излучение преобладает над спонтанным, испускаемый свет становится когерентным.

- Что такое p—n-переход? Как он устроен?
- Какие причины приводят к появлению электрического поля и потенциального барьера на границе p- и n-областей?
- Что такое основные и неосновные носители? Почему неосновные носители заряда сосредоточены вблизи p—n-перехода?
- Почему электрическое поле в p—n-переходе представляет собой барьер для основных носителей и крутой склон для неосновных?
- Как изменяется потенциальный барьер между p- и n-областями при приложении к нему прямого и обратного напряжения?
- Чем объясняется асимметричный вид вольт-амперной характеристики p—n-перехода?

- Нарисуйте схемы однополупериодного и двухполупериодного выпрямителей с использованием полупроводниковых диодов.
- Как устроен транзистор $p-n-p$-типа? Опишите, как он работает в схеме усилителя с общей базой.
- Рассмотрите работу в качестве усилителя транзистора $p-n-p$-типа. В какой полярности должны быть включены источники постоянного напряжения во входной и выходной цепях?
- Как оценить коэффициент усиления транзистора?
- Что такое интегральная микросхема?
- Поясните, почему полупроводниковые кристаллы поглощают и излучают свет в видимой и ближней инфракрасной области спектра?
- Какие особенности структуры энергетических зон кристаллов типа A^3B^5 определяют их применение в качестве материала для изготовления светодиодов и полупроводниковых лазеров?
- Какую роль играет закон сохранения импульса в процессах взаимодействия света с электронами и дырками в полупроводниках?
- Каким образом осуществляется инверсия заселенностей в полупроводниковых лазерах?

VIII. АТОМНОЕ ЯДРО И ЭЛЕМЕНТАРНЫЕ ЧАСТИЦЫ

§ 38. Строение атомного ядра

Существование атомных ядер впервые было экспериментально доказано в знаменитых опытах Резерфорда по рассеянию альфа-частиц. В этих опытах удалось также установить размеры ядра. Оказалось, что диаметр ядра имеет порядок 10^{-12}—10^{-13} м. В итоге возникла планетарная модель атома, которая была детально разработана Н. Бором. Теория Бора позволила объяснить многие наблюдаемые свойства атомов.

Атомное ядро на Земле и в космосе. Многочисленные опытные факты, такие, как естественная и искусственная радиоактивность, ядерные реакции, свидетельствуют о сложном строении ядра. Однако в окружающем нас земном мире атомные ядра, как правило, существуют только в своих основных энергетических состояниях. Большинство ядер ведут себя совершенно пассивно, выступая лишь носителями электрического заряда и массы, и никак не проявляют своих внутренних динамических свойств. Фактически на Земле все интересные ядерные явления происходят только в созданном руками человека искусственном мире ядерных реакторов и ускорителей заряженных частиц. Наиболее впечатляющие явления разыгрываются в гигантских ускорителях, способных сообщить разгоняемым частицам-«снарядам» огромные энергии, не встречающиеся в обычных условиях.

Иначе обстоит дело в масштабах Вселенной. Энергетические превращения, происходящие в недрах звезд, квазаров и других космических объектов, — это арена проявления динамических свойств ядер и элементарных частиц. В конечном итоге именно этим процессам мы обязаны всеми доступными на Земле источниками энергии. И сам состав окружающего нас сегодня материального мира представляет собой продукт ядерных реакций, происходящих на протяжении истории Вселенной.

Состав атомного ядра. По современным представлениям ядро атома любого элемента состоит из *протонов* и *нейтронов*, называемых *нуклонами*. Основные характеристики стабильных ядер — это *зарядовое число* Z, равное числу протонов, входящих в состав ядра, и *массовое число* A, равное полному числу нуклонов в ядре. Число N нейтронов в ядре, очевидно, равно разности $A - Z$.

Так как заряд протона представляет собой элементарный положительный заряд $e = 1{,}6 \cdot 10^{-19}$ Кл, то электрический заряд ядра равен Ze. В нейтральном атоме полное число электронов в электронной оболочке равно Z. Поэтому зарядовое число Z ядра совпадает с порядковым номером элемента в периодической системе Д. И. Менделеева и определяет все его химические свойства.

Наряду с термином «ядро атома» используется также термин *нуклид*. Нуклиды с одинаковыми зарядовыми числами Z, но различными числами нейтронов N, называются *изотопами*, так как соответствуют одному и тому же химическому элементу, т. е. одному и тому же месту в таблице Менделеева. Химические элементы имеют по нескольку изотопов и в природе встречаются в виде смесей определенного процентного состава. Нуклиды с одинаковыми массовыми числами A, но с различными Z и N, называются *изобарами* (т. е. одинаково тяжелыми).

Массы протонов и нейтронов очень близки: масса протона $m_p = 1836{,}15 m_e$, масса нейтрона $m_n = 1838{,}68 m_e$, где $m_e = 0{,}91095 \cdot 10^{-30}$ кг — масса электрона. Поэтому масса нуклида практически определяется общим числом A входящих в него нуклонов, а не значениями Z и N. За атомную единицу массы (а. е. м.) принимают $1/12$ часть массы нуклида изотопа углерода ${}^{12}_{6}\mathrm{C}$, содержащего 12 нуклонов. Поэтому в атомных единицах масса любого нуклона почти не отличается от единицы. В этих единицах масса ядра приближенно равна массовому числу A.

Энергия связи. Неточное совпадение массы нуклида с его массовым числом обусловлено не только различием масс протонов и нейтронов, но и тем, что их массы не складываются аддитивно в массу образуемого ими нуклида M:

$$M < Zm_p + Nm_n.$$

Разность между суммой масс протонов и нейтронов $Zm_p + Nm_n$ и массой ядра M называется *дефектом массы*. Дефект массы определяет *энергию связи* ядра $E_{св}$, т. е. ту энергию, которую необходимо затратить, чтобы расщепить ядро на отдельные нуклоны:

$$E_{св} = (Zm_p + Nm_n - M)c^2. \qquad (1)$$

Соотношение (1) является следствием общей релятивистской формулы $E_0 = m_0 c^2$, связывающей энергию покоя любого тела с его массой m_0. Очевидно, что энергия связи характеризует взаимодействие между нуклонами в ядре.

Ядерные силы. Силы, удерживающие нуклоны в ядре, называются *ядерными*. Эти силы представляют собой проявление самого интенсивного из известных в физике взаимодействий — так называемого

сильного взаимодействия. Ядерные силы, действующие между двумя протонами в ядре, примерно на два порядка больше кулоновских электростатических сил, действующих между ними, и в 10^{38} раз больше сил их гравитационного взаимодействия.

На основании опытных данных можно заключить, что нейтроны и протоны в ядре в отношении сильного взаимодействия ведут себя практически одинаково: ядерные силы между двумя протонами, двумя нейтронами или протоном и нейтроном неразличимы. Поэтому протоны и нейтроны в ядре рассматривают как два различных зарядовых состояния одной и той же частицы-нуклона. Независимость ядерных сил от зарядового состояния нуклонов называется *изотопической инвариантностью*.

Действие ядерных сил быстро спадает с расстоянием: на расстояниях, больших $2 \cdot 10^{-11}$ см, их действие не проявляется. Вплоть до расстояния порядка $0{,}7 \cdot 10^{-13}$ см они проявляются как силы притяжения, на меньших расстояниях — как силы отталкивания. Силы отталкивания настолько быстро растут с уменьшением расстояния, что нуклоны в ядре можно рассматривать как соприкасающиеся частицы неизменных размеров.

Размеры ядер. Размеры ядер зависят от числа содержащихся в них нуклонов. Средняя концентрация нуклонов в ядре для всех ядер с $A > 10$ практически одинакова. Это означает, что объем ядра примерно пропорционален числу нуклонов A и, следовательно, его радиус пропорционален $A^{1/3}$:

$$R = R_0 A^{1/3}, \tag{2}$$

где $R_0 \approx 1{,}2 \cdot 10^{-13}$ см. Плотность ядерной материи чрезвычайно велика по сравнению с плотностью обычных веществ и составляет около 10^{14} г/см^3. Такая плотность вещества характерна и для некоторых космических объектов, например нейтронных звезд — пульсаров.

Оказывается, что энергия связи $E_{\text{св}}$ также примерно пропорциональна числу A нуклонов в ядре, так что удельная энергия связи $E_{\text{св}}/A$ (т. е. энергия связи в расчете на один нуклон) слабо меняется при изменении A. Для большинства ядер значение $E_{\text{св}}/A$ лежит в интервале от 6 до 8 МэВ/нуклон.

Энергия связи и соотношение неопределенностей. Удельную энергию связи можно оценить на основе известных размеров ядра с помощью соотношения неопределенностей Гейзенберга. Когда нуклон находится внутри ядра, т. е. локализован в области размером порядка R, неопределенность Δp в значении его импульса составляет

$$\Delta p \sim h/R.$$

Так как само значение импульса p не может быть меньше этой неопределенности Δp, то эту же оценку можно использовать и для

импульса нуклона p. Соответствующее такому импульсу значение скорости нуклона массы $m \approx 1{,}7 \cdot 10^{-24}$ г составляет несколько десятых долей скорости света. Поэтому его кинетическая энергия определяется нерелятивистским выражением и равна

$$E_{\text{кин}} = \frac{p^2}{2m} \approx \frac{\hbar^2}{2mR^2} \approx 10 \text{ МэВ}.$$

Так как нуклон в ядре находится в связанном состоянии, то глубина потенциальной ямы, в которой он движется, имеет по крайней мере такой же порядок величины.

Полагая неопределенность в значении импульса одного порядка с самим импульсом, мы заведомо считали, что поведение нуклона в ядре нельзя описывать как движение классической частицы. К такому же выводу можно прийти на основе представлений о волнах де Бройля. Если оценить значение импульса нуклона на основе известной из опыта энергии связи на нуклон и подсчитать соответствующую такому импульсу длину волны де Бройля, то она окажется того же порядка величины, что и размер ядра.

Значение удельной энергии связи 10 МэВ/нуклон составляет менее одного процента энергии покоя нуклона $m_p c^2 \approx 1$ ГэВ. Поэтому действительно можно считать, что ядро состоит из отдельных нуклонов, которые сохраняют свою индивидуальность и внутри ядра.

Аналогичные соображения показывают, что в состав ядра не могут входить электроны. Если электрон локализован в ядре, т. е. в области размером $R \approx 10^{-13}$ см, то с помощью соотношений неопределенностей можно убедиться, что он будет ультрарелятивистским с кинетической энергией $E_{\text{кин}} \approx pc \approx 0{,}2$ ГэВ. Это значение существенно превосходит и энергию покоя электрона, равную 0,5 МэВ, и энергию связи ядра в расчете на одну частицу. Последнее, естественно, несовместимо с предположением, что электрон находится внутри ядра.

Капельная модель ядра. О приближенном постоянстве удельной энергии связи для различных ядер говорят как о насыщении ядерных сил. Оно фактически означает, что каждый нуклон эффективно взаимодействует не со всеми нуклонами ядра (в этом случае при $A \gg 1$ энергия связи была бы пропорциональна A^2), а лишь со своим ближайшим окружением. Эта ситуация до некоторой степени аналогична той, которая встречается при описании взаимодействия молекул жидкости. Такая аналогия в свое время послужила основой для создания так называемой *капельной модели* ядра, в которой принимается, что ядро ведет себя подобно капле несжимаемой заряженной жидкости. С помощью формулы (2) можно определить некоторые параметры такой ядерной жидкости. Для концентрации

нуклонов в ядре, очевидно, можно написать

$$n = \frac{A}{(4/3)\pi R^3} = \frac{3}{4\pi R_0^3} \approx 10^{38}\ \text{см}^{-3}.$$

Отсюда для плотности ядерного вещества имеем

$$\rho \approx m_p n \approx 1{,}4 \cdot 10^{14}\ \text{г/см}^3,$$

что совпадает с приведенным выше значением. Нетрудно оценить и среднее расстояние r между нуклонами в ядре:

$$r \approx n^{-1/3} \approx 2{,}3 \cdot 10^{-13}\ \text{см}.$$

Поскольку концентрация n нуклонов, плотность ρ вещества в ядре, а также среднее расстояние r между нуклонами практически одинаковы во всех ядрах, то ядерное вещество в капельной модели ядра можно считать практически несжимаемым.

Капельная модель позволила описать не только основное состояние ядра, но и некоторые из возбужденных состояний, рассматривая их как колебания формы поверхности капли. Однако эта чрезвычайно простая модель не в состоянии объяснить всего многообразия наблюдаемых свойств атомных ядер.

Кулоновское отталкивание протонов. Энергия связи нуклонов в ядре уменьшается из-за кулоновского отталкивания между протонами. Это кулоновское отталкивание является дальнодействующим в отличие от «контактного» сильного взаимодействия, действующего только между соприкасающимися нуклонами. Для легких ядер эффект кулоновского отталкивания не играет существенной роли, но для тяжелых ядер ситуация уже иная. В самом деле, энергия кулоновского отталкивания определяется попарным взаимодействием всех Z протонов ядра и потому пропорциональна $Z(Z-1)$, т. е. пропорциональна Z^2 при $Z \gg 1$. Энергия притяжения нуклонов из-за сильного взаимодействия, как уже отмечалось, пропорциональна полному числу нуклонов A. Так как числа протонов и нейтронов в устойчивых ядрах приблизительно одинаковы, то эта энергия фактически пропорциональна Z. Поэтому с ростом Z роль кулоновской энергии увеличивается. Этим объясняется уменьшение удельной энергии связи тяжелых ядер с возрастанием Z.

Ядерные связи между нуклонами наиболее прочны, когда числа протонов и нейтронов Z и N одинаковы, т. е. в ядре как бы образуются протон-нейтронные пары. Поэтому у легких стабильных ядер, где роль кулоновского взаимодействия невелика, числа протонов и нейтронов одинаковы. Однако у ядер с большими атомными номерами для обеспечения устойчивости требуются дополнительные нейтроны. Это обусловлено возрастанием относительной

роли кулоновского взаимодействия между протонами. Устойчивость ядра с ростом Z достигается вплетением все большего числа нейтронов. Это отчетливо видно из диаграммы устойчивых изотопов, показанной на рис. 124. У ядер элементов, следующих за свинцом ($Z > 82$), уже так много протонов, что полная их стабильность оказывается вообще невозможной.

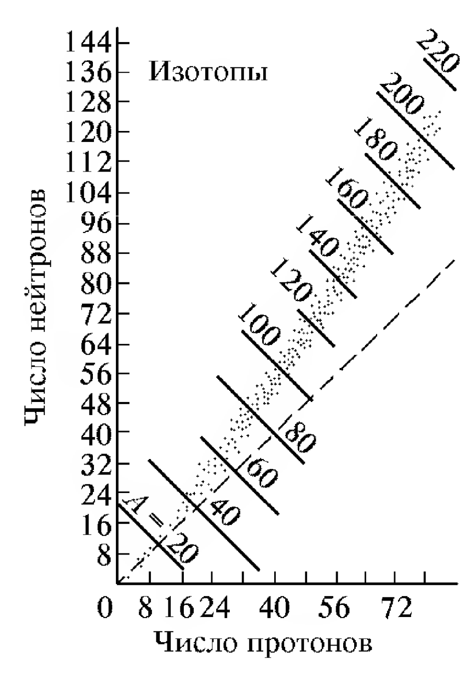


Рис. 124. Число протонов и нейтронов в стабильных ядрах

Наибольшей устойчивостью и распространенностью в природе отличаются ядра, у которых число протонов или число нейтронов равно одному из так называемых *магических чисел*: 2, 8, 20, 28, 50, 82, 126. Если у ядра одновременно являются магическими как число нейтронов, так и число протонов, то такие дважды магические ядра отличаются особенно большой устойчивостью. Таких ядер всего пять: ${}^{4}_{2}He$, ${}^{16}_{8}O$, ${}^{40}_{20}Ca$, ${}^{48}_{20}Ca$ и ${}^{208}_{82}Pb$. Повышенная устойчивость магических ядер объясняется так называемой *оболочечной моделью* ядра.

- Как проявляет себя атомное ядро в земных условиях и в масштабах Вселенной?
- Как связаны между собой порядковый номер элемента в периодической системе с зарядом его ядра?
- Что такое изотопы и изобары?
- Почему масса ядра не равна сумме масс образующих его протонов и нейтронов?
- Как с помощью соотношений неопределенностей оценить энергию связи нуклонов в ядре?
- Почему можно считать, что входящие в состав атомного ядра протоны и нейтроны сохраняют свою индивидуальность и внутри ядра? Почему в состав ядра не могут входить электроны?
- Что такое капельная модель ядра?
- Почему относительное число нейтронов в ядре больше у тяжелых ядер?
- Какие атомные ядра отличаются наибольшей устойчивостью?

§ 39. Радиоактивность. Ядерные реакции

Наряду со стабильными ядрами существуют радиоактивные ядра, в которых происходит самопроизвольное изменение состава. Большая

часть известных радиоактивных ядер получена искусственно путем бомбардировки мишеней различными частицами. Известно несколько видов радиоактивного распада.

Альфа-распад. При α-распаде из ядра спонтанно вылетает альфа-частица — ядро атома гелия ^4_2He. При этом зарядовое число Z ядра в соответствии с законом сохранения электрического заряда уменьшается на два и образуется ядро нового химического элемента, который сдвинут относительно исходного на две клетки периодической системы.

Чтобы α-распад был возможен, нужно, чтобы энергия связи исходного (материнского) ядра $E_{св}(Z, A)$ была меньше суммы энергии связи образующегося (дочернего) ядра $E_{св}(Z-2, A-4)$ и энергии связи альфа-частицы $E_{св}(2, 4)$. Кинетическая энергия $E_{кин}$ вылетающей альфа-частицы определяется законом сохранения энергии

$$E_{кин} = E_{св}(Z-2, A-4) + E_{св}(2, 4) - E_{св}(Z, A). \qquad (1)$$

В соотношении (1) не учтена кинетическая энергия отдачи ядра, потому что масса дочернего ядра всегда много больше массы альфа-частицы. При этом, разумеется, считается, что материнское ядро неподвижно. Несколько иначе условие возможности альфа-распада можно сформулировать такими словами: масса (покоя) материнского ядра должна быть больше суммы масс покоя дочернего ядра и вылетающей альфа-частицы.

В простейшей теории α-распада, предложенной Г. Гамовым, предполагается, что альфа-частица уже существует внутри материнского ядра, которое является для нее своего рода потенциальной ямой. Энергия альфа-частицы в ядре недостаточна для преодоления ограничивающего эту яму потенциального барьера. Вылет альфа-частицы из ядра оказывается возможным благодаря квантовомеханическому туннельному эффекту, заключающемуся в отличной от нуля вероятности прохождения частицы под потенциальным барьером конечных размеров. Момент вылета альфа-частицы из ядра непредсказуем, поскольку явление радиоактивного распада имеет вероятностный характер. Квантовая теория позволяет найти лишь вероятность распада за единицу времени или обратную ей величину, называемую средним временем жизни радиоактивного ядра.

Бета-распад. При β-распаде из ядра вылетают электрон и электронное антинейтрино. Существование этой электрически нейтральной частицы было предположено Паули для объяснения кажущегося нарушения закона сохранения энергии в элементарном акте β-распада. В β-распаде распределение энергии между вылетающим электроном и антинейтрино имеет случайный характер. Поэтому в отличие от альфа-частиц, вылетающих из данного ядра с вполне определенной энергией, вылетающие электроны могут иметь разную энергию.

При β-распаде вылетающий электрон не существует внутри ядра, а образуется там при превращении нейтрона в протон. В этом смысле говорят, что β-распад — это не внутриядерный, а внутринуклонный процесс. Он затрагивает более глубокие изменения структуры вещества, чем α-распад. Теория β-распада была разработана Ферми на основе предположения о так называемом слабом взаимодействии, описывающем превращение нейтрона в протон.

Нестабильность нейтрона. Такой процесс происходит не только внутри ядра, но и со свободным нейтроном, среднее время жизни которого составляет около 15 минут. При этом нейтрон ${}^1_0\mathrm{n}$ распадается на протон ${}^1_1\mathrm{p}$, электрон ${}^{\;0}_{-1}\mathrm{e}$ и антинейтрино $\bar{\nu}$:

$$ {}^1_0\mathrm{n} \rightarrow {}^1_1\mathrm{p} + {}^{\;0}_{-1}\mathrm{e} + \bar{\nu}. \qquad (2)$$

При β-распаде массовое число A ядра не меняется, а зарядовое число Z увеличивается на единицу: образуется новый химический элемент, который сдвинут в периодической системе вправо на одну клетку.

Наряду с электронным β-распадом существует позитронный β^+-распад, при котором из ядра вылетает позитрон ${}^{\;0}_{+1}\mathrm{e}$ и электронное нейтрино ν. При позитронной радиоактивности один из протонов ядра превращается в нейтрон с испусканием позитрона и нейтрино. В периодической системе происходит сдвиг влево на одну клетку без изменения массового числа. Для свободного протона в обычных условиях такой распад невозможен, так как его масса меньше массы нейтрона.

Гамма-распад. В отличие от α- и β-радиоактивных распадов так называемая γ-радиоактивность ядер не связана с изменением внутренней структуры ядра и не сопровождается изменением зарядового или массового чисел. Вылет из ядра γ-кванта (фотона высокой энергии) происходит при спонтанном переходе ядра из некоторого долгоживущего возбужденного состояния в основное состояние.

Закон радиоактивного распада. Радиоактивный распад не зависит от внешних условий в широком интервале изменения таких параметров, как температура и давление. Закон радиоактивного распада, т. е. зависимость от времени числа $N(t)$ еще не распавшихся к данному моменту ядер некоторого радиоактивного образца, легко получить из предположения, что вероятность распада является постоянной для данного вида ядер величиной, не зависящей от того, сколько времени уже «прожило» данное радиоактивное ядро. За малый промежуток времени Δt количество нераспавшихся ядер $N(t)$ изменится на число ΔN, пропорциональное количеству имеющихся ядер $N(t)$ и этому промежутку Δt:

$$\Delta N = -\lambda N(t)\Delta t. \qquad (3)$$

Коэффициент пропорциональности λ — это и есть не зависящая от времени вероятность распада ядра. Знак минус в (3) соответствует уменьшению со временем числа нераспавшихся ядер. Формула (3) означает, что скорость dN/dt изменения искомой функции $N(t)$ пропорциональна самой функции:

$$\frac{dN}{dt} = -\lambda N.$$

Отсюда следует, что $N(t)$ убывает со временем по экспоненциальному закону:

$$N(t) = N_0 e^{-\lambda t}, \tag{4}$$

где N_0 — начальное число радиоактивных ядер при $t = 0$. Можно показать, что величина $\tau = 1/\lambda$ — это среднее время жизни радиоактивного ядра.

Часто закон радиоактивного распада (4) записывают в виде

$$N(t) = N_0 \cdot 2^{-t/T}, \tag{5}$$

используя основание 2 вместо e. В этом случае величина T называется *периодом полураспада*. Очевидно, что T — это время, в течение которого распадается половина первоначального количества ядер. Сравнивая правые части формул (4) и (5), видим, что

$$T = \frac{1}{\lambda} \ln 2 = \tau \ln 2 \approx 0{,}71\tau. \tag{6}$$

Возраст Земли. Закон радиоактивного распада позволяет оценить возраст Земли на основе известного процентного содержания в урановой руде радиоактивных изотопов урана ${}^{238}_{92}U$ и ${}^{235}_{92}U$. Измеренные времена жизни этих изотопов составляют соответственно $\tau_1 = 6{,}52 \cdot 10^9$ лет и $\tau_2 = 1{,}02 \cdot 10^9$ лет. Их процентное содержание в естественной смеси составляет 99,3 % для ${}^{238}_{92}U$ и 0,7 % для ${}^{235}_{92}U$. Естественно предположить, что в момент образования химических элементов, в том числе урана, содержание этих изотопов было одинаковым. Поэтому в выражении (4), записанном для каждого изотопа урана, следует взять одинаковые значения N_0 и соответствующие имеющемуся в настоящее время процентному составу значения N_{235} и N_{238}:

$$N_{238} = N_0 e^{-t/\tau_1}, \qquad N_{235} = N_0 e^{-t/\tau_2},$$

откуда

$$\frac{N_{235}}{N_{238} + N_{235}} \approx \frac{N_{235}}{N_{238}} = \frac{e^{-t/\tau_1}}{e^{-t/\tau_2}} = 0{,}007.$$

Это дает для t значение $6 \cdot 10^9$ лет, что хорошо согласуется с другими оценками возраста Земли или Солнечной системы, полученными на основе геологических и космологических соображений.

Ядерные реакции. *Ядерная реакция* — это взаимодействие атомного ядра с другим ядром или элементарной частицей, в результате

которого происходит изменение его состава и структуры. Из-за короткодействующего характера ядерных сил ядерные реакции происходят при сближении частиц на расстояния порядка 10^{-13} см.

При ядерных реакциях выполняются законы сохранения энергии, импульса, момента импульса, электрического заряда, числа нуклонов (барионного заряда). Кроме того, выполняется закон сохранения так называемого лептонного заряда, согласно которому рождение электрона обязательно сопровождается рождением электронного антинейтрино, а рождение позитрона рождением электронного нейтрино. Образование электрон—позитронной пары может происходить и без появления нейтрино и антинейтрино.

Первая искусственная ядерная реакция была осуществлена Резерфордом при облучении азота альфа-частицами:

$$^{14}_{7}\mathrm{N} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{17}_{8}\mathrm{O} + {}^{1}_{1}\mathrm{H}.$$

При ядерных реакциях могут образовываться новые радиоактивные изотопы, которых нет на Земле в естественном состоянии. Радиоактивность ядра, возникающая в результате ядерной реакции, называется *искусственной*. Она была открыта на ядерной реакции

$$^{27}_{13}\mathrm{Al} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{30}_{15}\mathrm{P}^* + {}^{1}_{0}\mathrm{n}.$$

Радиоактивный изотоп фосфора $^{30}_{15}P^*$ представляет собой источник β^+-излучения. Его ядро испускает позитрон и электронное нейтрино:

$$^{30}_{15}P^* \rightarrow {}^{30}_{14}\mathrm{Si} + {}^{0}_{+1}\mathrm{e} + \nu.$$

Существование нейтрона было обнаружено до открытия искусственной радиоактивности. Нейтрон был открыт на ядерной реакции

$$^{9}_{4}\mathrm{Be} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{12}_{6}\mathrm{C} + {}^{1}_{0}\mathrm{n}.$$

Создание ускорителей заряженных частиц значительно расширило возможности осуществления различных ядерных реакций. Теперь в качестве снарядов для бомбардировки ядер-мишеней можно использовать не только альфа-частицы, испускаемые радиоактивными препаратами, но и разогнанные до высокой энергии протоны, дейтроны и электроны. Первое расщепление ядра, выполненное с помощью искусственно разогнанных протонов, было осуществлено на реакции

$$^{7}_{3}\mathrm{Li} + {}^{1}_{1}\mathrm{P} \rightarrow {}^{4}_{2}\mathrm{He} + {}^{4}_{2}\mathrm{He}.$$

Ядерные реакции могут происходить также и под действием гамма-квантов достаточно больших энергий — так называемые *фотоядерные реакции*. Они становятся возможными, когда энергия гамма-кванта превышает энергию отделения нуклонов от ядра.

Энергетические превращения при ядерных реакциях. Ядерные реакции сопровождаются энергетическими превращениями. Понять

характер этих превращений можно с помощью графика зависимости удельной энергии связи нуклонов в ядре от массового числа ядра (рис. 125). Экспериментальное значение энергии в расчете на один нуклон составляет около 8 МэВ/нуклон. Сначала с ростом массового числа A она увеличивается от 1,1 МэВ/нуклон у ядра дейтерия ${}^{2}_{1}D$ до 8,8 МэВ/нуклон у изотопа железа ${}^{56}_{26}Fe$, а затем постепенно убывает до 7,6 МэВ/нуклон у изотопа урана ${}^{238}_{92}U$. Такое поведение удельной энергии связи объясняется неполным насыщением ядерных сил у легких ядер, когда число соседей у каждого нуклона меньше, чем это возможно при их плотной упаковке, и возрастанием роли кулоновского отталкивания протонов в ядрах с большим массовым числом.



Рис. 125. Зависимость удельной энергии ядер от массового числа

Характер зависимости удельной энергии связи от массового числа объясняет выделение энергии как при делении тяжелых ядер на осколки, лежащие в середине периодической системы элементов, так и при синтезе легких ядер. Поэтому возможны два принципиально различных способа освобождения ядерной энергии для практического применения.

Деление тяжелых ядер. Первый путь связан с использованием реакции деления, которая может быть осуществлена для элементов от трансурановых до железа. Практически для этой цели используется уран. При облучении ${}^{235}_{92}U$ нейтронами его ядро захватывает нейтрон и делится на два радиоактивных ядра-осколка с массовыми числами порядка 100, для которых удельная энергия связи близка к максимальной. При этом в каждом элементарном акте распада ядра урана образуется от двух до четырех нейтронов. Например,

$$ {}^{235}_{92}U + {}^{1}_{0}n \rightarrow {}^{236}_{92}U^* \rightarrow {}^{144}_{56}Ba^* + {}^{89}_{36}Kr^* + 3\,{}^{1}_{0}n, $$

$$ {}^{235}_{92}U + {}^{1}_{0}n \rightarrow {}^{236}_{92}U^* \rightarrow {}^{140}_{54}Xe^* + {}^{64}_{38}Sr^* + 2\,{}^{1}_{0}n. $$

Поскольку в каждом акте испускается нейтронов больше, чем поглощается, то при делении ${}^{235}_{92}U$ может возникнуть самоподдерживающаяся *цепная реакция*, т. е. процесс, в котором определенная ядерная реакция вызывает последующие реакции такого же типа.

Для осуществления цепной реакции необходимо, чтобы так называемый коэффициент размножения нейтронов — отношение числа нейтронов в каком-либо поколении к числу нейтронов предшествую-

щего поколения — был больше единицы. Коэффициент размножения определяется не только числом образующихся в каждом элементарном акте нейтронов, но и условиями, в которых протекает реакция: часть образующихся нейтронов поглощается посторонними ядрами или уходит за пределы зоны реакции. Минимальная масса урана, при которой возможна цепная реакция, называется *критической массой*.

Преобладающей в природной смеси изотоп $^{238}_{92}U$ также может захватывать нейтроны. При этом образуется радиоактивный изотоп нептуния:

$$^{238}_{92}U + ^{1}_{0}n \rightarrow ^{239}_{92}U^* \underset{23\text{ мин}}{\longrightarrow} ^{239}_{93}Np^* + ^{0}_{-1}e + \nu,$$

который в свою очередь в результате β-распада превращается в плутоний:

$$^{239}_{93}Np^* \underset{2,3\text{ дня}}{\longrightarrow} ^{239}_{94}Pu + ^{0}_{-1}e + \nu.$$

Плутоний относительно стабилен (период полураспада 24000 лет) и также может использоваться как ядерное топливо в ядерных реакциях.

Устройство, в котором осуществляется управляемая ядерная цепная реакция, сопровождающаяся выделением энергии, называется *ядерным реактором*. Первый ядерный реактор построен в декабре 1942 г. в США под руководством Э. Ферми.

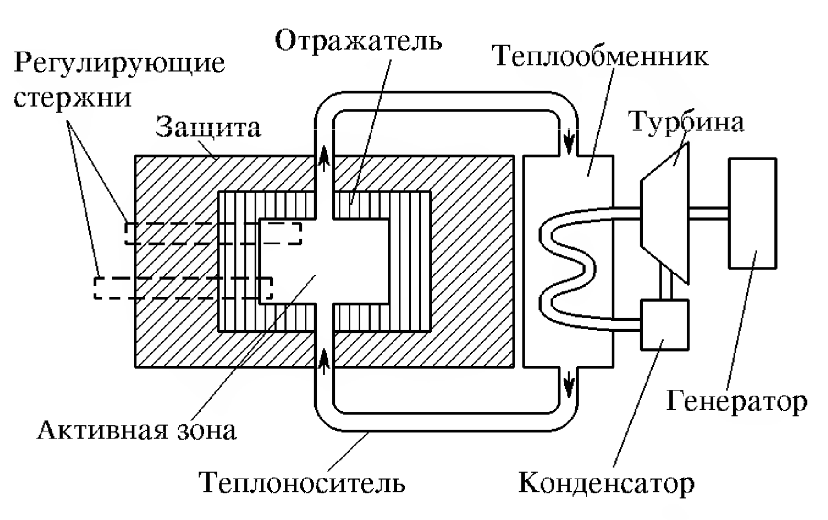


Рис. 126. Схема энергетического ядерного реактора

В любом ядерном реакторе имеются следующие основные части (рис. 126): активная зона, где находится ядерное топливо, протекает цепная реакция деления ядер и выделяется энергия; отражатель нейтронов, окружающий активную зону; теплоноситель, отводящий выделяющуюся энергию; система регулирования цепной реакции; радиационная защита.

Об экологических проблемах ядерной энергетики. Большинство образующихся в реакциях деления ядер-осколков радиоактивны

и дают начало целым цепочкам последовательных превращений типа

$$^{90}_{36}\mathrm{Kr}^* \underset{33\ \text{с}}{\rightarrow} {}^{90}_{37}\mathrm{Rb}^* + {}^{0}_{-1}e + \nu;$$

$$^{90}_{37}\mathrm{Rb}^* \underset{274\ \text{с}}{\rightarrow} {}^{90}_{38}\mathrm{Sr}^* + {}^{0}_{-1}e + \nu;$$

$$^{90}_{38}\mathrm{Sr}^* \underset{28\ \text{лет}}{\rightarrow} {}^{90}_{39}\mathrm{Y} + {}^{0}_{-1}e + \nu.$$

Образование большого числа радиоактивных нуклидов ставит серьезные экологические проблемы при развитии ядерной энергетики.

Реакции синтеза. Второй путь освобождения ядерной энергии связан с реакциями синтеза. Чтобы осуществить реакцию синтеза, нужно сблизить ядра до расстояний 10^{-12} см, когда проявляется действие ядерных сил. При этом приходится преодолевать кулоновское отталкивание ядер, для чего их необходимо разогнать до очень больших скоростей. Реакции ядерного синтеза становятся возможными при температурах порядка 10^8 К. Выделяющаяся при слиянии легких ядер энергия в расчете на один нуклон еще больше, чем при распаде тяжелых ядер. Например:

$$^{2}_{1}\mathrm{D} + {}^{2}_{1}\mathrm{D} \rightarrow {}^{3}_{1}\mathrm{T} + {}^{1}_{1}\mathrm{H} + 4{,}0\ \text{МэВ};$$

$$^{2}_{1}\mathrm{D} + {}^{2}_{1}\mathrm{D} \rightarrow {}^{3}_{2}\mathrm{He} + {}^{1}_{0}n + 3{,}25\ \text{МэВ};$$

$$^{2}_{1}\mathrm{D} + {}^{3}_{1}\mathrm{T} \rightarrow {}^{4}_{2}\mathrm{He} + {}^{1}_{0}n + 17{,}7\ \text{МэВ}.$$

Получение высоких температур и удержание нагретой до миллиарда градусов плазмы представляет собой труднейшую научно-техническую задачу на пути осуществления управляемого термоядерного синтеза. Тем не менее, осуществление управляемой реакции синтеза очень заманчиво благодаря экологической чистоте процессов и практически неисчерпаемым запасам термоядерного топлива.

- Поясните, почему предположение о существовании α-частицы внутри ядра не противоречит основным принципам квантовой физики.
- Как найти кинетическую энергию вылетающей из радиоактивного ядра α-частицы?
- Почему β-распад следует считать внутринуклонным, а не внутриядерным процессом, как α-распад?
- Почему вылетающие из ядра при β-распаде электроны могут иметь разные энергии?

- Почему в обычных условиях нейтрон неустойчив и распадается на протон, электрон и антинейтрино, а протон — стабилен?
- Как можно на опыте разделить разные виды радиоактивных излучений (α, β, γ) из одного радиоактивного препарата?
- Как на основе закона радиоактивного распада оценить возраст Земли?
- Возраст Земли — несколько миллиардов лет, а период полураспада радия — 1600 лет. Почему же на Земле еще сохранилось какое-то количество радия?
- Что такое искусственная радиоактивность? С чем связана высокая эффективность облучения нейтронами для стимулирования ядерных реакций?
- Почему выделение внутриядерной энергии возможно как при реакциях деления тяжелых ядер, так и при синтезе легких ядер?
- В чем суть экологических проблем, возникающих в связи с развитием ядерной энергетики?
- С чем связаны трудности на пути осуществления управляемой реакции термоядерного синтеза?

§ 40. Элементарные частицы

В современной физике термин *элементарные частицы* употребляется не в своем обычном значении чего-то первичного, неразложимого на более простое, а для наименования большой группы мельчайших субъядерных частиц. В эту группу входят протон, нейтрон, электрон, фотон, π-мезон, мюон, нейтрино нескольких типов, так называемые странные частицы (K-мезоны, гипероны), очарованные частицы, промежуточные векторные бозоны и т. д. — всего к настоящему времени известно более 350 частиц, в основном нестабильных.

Большинство перечисленных частиц не удовлетворяют обычному определению элементарности, поскольку по современным представлениям они сами являются составными системами. Объединяющий их признак заключается в том, что они представляют форму материи, не ассоциированной в ядра и атомы.

Превращения элементарных частиц. Наиболее важное свойство всех элементарных частиц — способность к взаимным превращениям, т. е. способность рождаться и уничтожаться (испускаться и поглощаться). Все процессы с элементарными частицами, включая их распады, протекают через последовательность актов поглощения и испускания, в которых непременно выполняются законы сохранения.

Фундаментальные взаимодействия. Процессы с участием различных элементарных частиц сильно различаются по интенсивности

протекания, т. е. по характерным временам и энергиям. В соответствии с этим взаимодействия, в которых они участвуют, феноменологически подразделяют на сильное, электромагнитное, слабое и гравитационное. Сильное взаимодействие приводит к наиболее прочной связи элементарных частиц; именно оно обусловливает связь протонов и нейтронов в атомных ядрах.

Электромагнитное взаимодействие ответственно за связь атомных электронов с ядрами и связь атомов в молекулах и конденсированных средах. Между элементарными частицами это взаимодействие осуществляется через электромагнитное поле; для его существования наличие электрического заряда у частицы не обязательно. Например, не обладающий элементарным зарядом нейтрон имеет магнитный момент и участвует в электромагнитном взаимодействии.

Слабое взаимодействие проявляется в сравнительно медленно протекающих процессах распада некоторых элементарных частиц и атомных ядер. Например, благодаря слабому взаимодействию свободный нейтрон распадается на протон, электрон и электронное антинейтрино. Несмотря на сравнительно малую интенсивность и короткодействие, слабое взаимодействие играет очень важную роль в устройстве Вселенной. Например, если бы удалось «выключить» слабое взаимодействие, то погасло бы Солнце, так как был бы невозможен процесс превращения протона в нейтрон, позитрон и нейтрино, в результате которого четыре протона в конечном счете синтезируются в ядро $^4_2\mathrm{He}$. Этот процесс служит источником энергии Солнца и большинства звезд.

Все без исключения частицы участвуют в гравитационном взаимодействии, которое, однако, на субатомных расстояниях порядка 10^{-13} см и меньше не играет практически никакой роли.

В зависимости от способности к участию в тех или иных видах взаимодействий все элементарные частицы, за исключением фотона, разбиваются на две основные группы: адроны и лептоны. Адроны наряду с электромагнитным и слабым взаимодействием участвуют в сильном взаимодействии. Лептоны участвуют только в электромагнитном и слабом взаимодействиях.

Поиски единого взаимодействия. Относительная роль разных видов взаимодействий (сильного, электромагнитного, слабого) в процессах с элементарными частицами зависит от энергии частиц. Поэтому деление взаимодействий на виды в зависимости от интенсивности процессов надежно осуществляется только при не слишком высоких энергиях. В современной физике растет уверенность, что все взаимодействия в природе тесно связаны между собой и по существу являются различными проявлениями некоторого единого по-

ля. Объединение всех взаимодействий остается пока нерешенной задачей физической теории.

Очень важным этапом на этом пути стал успех в создании теории, объединившей электромагнитное и слабое взаимодействия в единое так называемое электрослабое взаимодействие. В этой новой теории электромагнитное поле и поле слабого взаимодействия рассматриваются как разные компоненты одного поля, в котором наряду с квантом электромагнитного поля фотоном участвуют три новых частицы — промежуточные векторные бозоны W^+, W^-, Z^0. Необходимость существования этих частиц вытекает в новой теории из аналогии с электромагнитным взаимодействием, где электрический заряд играет двойную роль: с одной стороны, он является сохраняющейся величиной, определяющей само существование взаимодействия, а с другой стороны — источником электромагнитного поля, кванты которого осуществляют взаимодействие между заряженными частицами. Будучи аналогичными фотону, новые частицы все же резко отличаются от него наличием заряда и массы.

Электрический заряд у W^+-бозона такой же, как и у протона, у W^- — как у электрона, а Z^0-бозон электрически нейтрален. Новая теория не только предсказала существование этих частиц, но и позволила выразить их массы через известные из опыта константы, характеризующие электромагнитное и слабое взаимодействия при низких энергиях. Они оказались равными $m_{W^+} = m_{W^-} \approx 81$ ГэВ и $m_Z \approx 92$ ГэВ. Большая масса этих частиц (примерно в 100 раз больше, чем у протона) не давала возможности обнаружить их и исследовать экспериментально, пока не было достаточно мощных ускорителей. В 1983 г. в Женеве на ускорителе, где происходили столкновения протонов и антипротонов во встречных пучках с энергией 270 ГэВ в каждом пучке, все три частицы были открыты. Их массы оказались в блестящем согласии с теоретически предсказанными значениями.

О механизме фундаментальных взаимодействий. Идея о том, что в мире элементарных частиц взаимодействие осуществляется посредством обмсна квантами какого-либо поля, родилась в физикс сщс в 30-х годах XX века, когда Юкава предположил, что сильное взаимодействие между нуклонами обязано своим происхождением гипотетическим частицам, получившим название мезонов. Массу этих частиц можно оценить с помощью соотношения неопределенностей. Неопределенность ΔE значения энергии ядра при испускании некоторой частицы (мезона) массы m порядка энергии этой частицы: $\Delta E = mc^2$. Эта неопределенность в значении энергии существует в течение времени Δt пролета мезона внутри ядра, которое дается отношением размера ядра R к скорости частицы $v = p/m$: $\Delta t \approx Rm/p$. Входящее сюда

значение импульса p оценим из соотношения неопределенностей, учитывая, что мезон локализован внутри ядра: $p \approx h/R$. Отсюда следует, что $\Delta t \approx mR^2/h$. Подставляя ΔE и Δt в соотношение $\Delta E \cdot \Delta t \approx h$, получим оценку для массы мезона m:

$$m \approx \frac{h}{Rc} \approx 3 \cdot 10^{-25}\ \text{г} \approx 300 m_e \approx 150\ \text{МэВ}. \qquad (1)$$

Частицы с такой массой были впоследствии открыты на опыте и получили название π-мезонов (пионов). Существуют три вида пионов: π^+, π^- и π^0, массы которых составляют $m_{\pi^+} = m_{\pi^-} \approx 140$ МэВ, $m_{\pi^0} \approx 135$ МэВ.

Радиус фундаментальных взаимодействий. Описанная выше идея о том, что физический механизм взаимодействия заключается в обмене квантами некоторого поля, оказалась очень плодотворной. Найденная связь между радиусом действия сил и массой обмениваемых частиц имеет универсальный характер: для любых видов взаимодействия радиус сил, как видно из (1), обратно пропорционален массе частицы:

$$R \approx \frac{h}{mc}. \qquad (2)$$

Электромагнитное взаимодействие характеризуется бесконечным радиусом действия сил, и поэтому масса фотона равна нулю. Как мы видели, для ядерных сил с радиусом действия 10^{-13} см масса мезона оказалась около 150 МэВ. В случае слабого взаимодействия радиус действия сил составляет 10^{-16} см, что дает для массы векторных бозонов приведенное выше значение.

Обсуждаемая связь между массой обмениваемой частицы и радиусом соответствующего взаимодействия по существу отражает корпускулярно-волновой дуализм, присущий всем квантовым объектам. Действительно, кванту поля с энергией $E = h\nu$, рассматриваемому как частица, следует сопоставить выражение $E = mc^2$, откуда для радиуса взаимодействия имеем

$$R \approx \lambda = \frac{c}{\nu} = \frac{h}{mc}. \qquad (3)$$

Величину $h/(mc)$ называют *комптоновской длиной волны* частицы с массой m. Именно она характеризует пространственный масштаб взаимодействия, осуществляемого обменом частицами такой массы.

Свойства электрослабого взаимодействия. Большая масса промежуточных векторных бозонов и связанный с этим очень малый радиус слабого взаимодействия приводят к важным для нас следстви-

ям. Важнейшая вытекающая отсюда особенность заключается в том, что при низких энергиях обусловленные слабым взаимодействием процессы протекают очень медленно. Даже при температурах и плотностях, которые господствуют в центре Солнца, обусловленные слабым взаимодействием процессы синтеза протонов в легкие ядра приводят к скорости выделения теплоты на единицу массы, приблизительно в 100 раз меньшей, чем при естественном обмене веществ в организме человека. Медленность этих процессов обеспечивает постепенное «выгорание» Солнца на данном этапе его эволюции, когда оно своим излучением дает жизнь всему земному.

При высоких энергиях, достижимых на современных ускорителях с встречными протон-антипротонными пучками, ситуация кардинально изменяется. Обмен тяжелыми векторными бозонами происходит теперь столь же эффективно, как и обмен фотонами. При очень высоких энергиях слабое взаимодействие может стать даже сильнее электромагнитного. Здесь законы обычной электродинамики уже не работают. Правильное описание процессов дает новая теория — теория электрослабого взаимодействия, учитывающая как обмен фотонами, так и обмен промежуточными векторными бозонами. Эта теория продолжает теорию Максвелла в область малых расстояний или, что то же самое, в область высоких энергий.

Сильное взаимодействие. Кварки. Значительные успехи достигнуты в последние годы и в теории сильного взаимодействия, где нуклоны уже перестали играть роль исходных частиц. Оказалось, что все адроны можно представить как составные частицы, образованные из так называемых *кварков*.

Кварки бывают шести сортов. В строении окружающего нас мира наиболее важны u- и d-кварки, из которых построены протоны и нейтроны. Кварки обладают дробным электрическим зарядом. Заряд u-кварка составляет $+2/3$ элементарного заряда, заряд d-кварка $-1/3$. Протон состоит из двух u- и одного d-кварка. Нейтрон состоит из одного u- и двух d-кварков. Когда d-кварк внутри нейтрона превращается в u-кварк, заряд нуклона изменяется на $(+2/3) - (-1/3) = 1$, что воспринимается наблюдателем как превращение нейтрона в протон.

Кварки не наблюдаются в свободном состоянии и по современным представлениям принципиально не могут существовать сами по себе вне адронов. Всякая попытка освободить кварк из «заключения» в адроне неизбежно заканчивается рождением пары кварк-антикварк, образующих в совокупности мезон. Взаимодействие между кварками в адроне осуществляется посредством обмена так называемыми *глюонами* — гипотетическими электрически нейтральными частицами с нулевой массой покоя. Основной вклад в массы протонов и нейтронов

дают не кварки, из которых они состоят, а сильное взаимодействие между кварками, обусловленное глюонами. В этом смысле говорят, что глюоны переносят массу, хотя сами ею не обладают.

Состояние кварков характеризуют рядом квантовых чисел, получивших экзотические названия, одно из которых — цвет дало название всей науке, описывающей взаимодействие кварков с глюонами, — *квантовая хромодинамика.* Гипотеза кварковой структуры оказалась необходимой для понимания динамики различных процессов с участием адронов. Кроме того, эта гипотеза позволила объяснить эмпирически установленную классификацию адронов.

Внутренние симметрии. Руководящая идея в развитии теории элементарных частиц основана на представлении о внутренних симметриях. Например, сильное взаимодействие симметрично относительно поворотов в абстрактном изотопическом пространстве. Одним из проявлений этой симметрии является зарядовая независимость ядерных сил. Так называемая калибровочная симметрия отвечает тому факту, что некоторые сохраняющиеся величины, называемые «зарядами» (например, электрический заряд), являются одновременно источниками полей, переносящих взаимодействия между частицами, обладающими данным типом «заряда». С каждым типом симметрии в физике связан определенный закон сохранения.

Аннигиляция частицы и античастицы. Соображения симметрии приводят к неизбежному выводу о том, что у каждой элементарной частицы существует «двойник» — *античастица*, которая отличается от частицы только знаком некоторых характеристик взаимодействий (например, электрического заряда, магнитного момента, лептонного и барионного зарядов). У некоторых частиц, в частности у фотона, античастица совпадает с самой частицей. Такие частицы называют истинно нейтральными.

При встрече частицы со своей античастицей происходит их *аннигиляция.* Например, при аннигиляции электрона и позитрона они превращаются в два, три или несколько γ-квантов. Один γ-квант излучиться не может, так как это несовместимо с законами сохранения. При аннигиляции тяжелых частиц и античастиц возникают не столько γ-кванты, сколько другие легкие частицы, например π-мезоны при аннигиляции протона и антипротона. Наряду с аннигиляцией, при достаточно большой энергии возможен и обратный процесс *рождения пары* частица-античастица.

Великое объединение. Значительные усилия прилагаются в настоящее время в попытках рассмотреть на единой основе не только электромагнитное и слабое, но и сильное взаимодействие. Наблюдаемые большие различия между этими взаимодействиями счи-

таются обусловленными нарушением симметрии при доступных в настоящее время энергиях. Их единая природа может проявиться только при энергии частиц во встречных пучках порядка 10^{14} ГэВ. При этом кварки и лептоны окажутся однотипными объектами и станут возможными их взаимные превращения. Следствием таких представлений явилось предсказание нестабильности свободного протона со средним временем жизни 10^{30}—10^{32} лет, что существенно превышает возраст Вселенной. Эта теория известна под названием Великого объединения. Теория, которая сумеет включить и гравитацию, уже заранее получила название Суперобъединение.

Теории Великого объединения актуальны лишь при столь высоких энергиях, какие могли быть только на самых ранних этапах существования Вселенной. Таким образом, физика элементарных частиц, прорываясь в область высоких энергий, соединилась с современной космологией — теорией эволюции Вселенной. Появилась новая наука — космомикрофизика.

- Какой смысл вкладывается в понятие «элементарная частица» в современной физике?
- Укажите роль каждого из видов фундаментальных взаимодействий в различных физических явлениях.
- Почему слабое взаимодействие, несмотря на свою малую интенсивность и короткодействующий характер, играет очень важную роль в эволюции Вселенной?
- Как с помощью соотношения неопределенностей оценить массу мезона, переносящего сильное взаимодействие?
- Каковы современные физические представления о механизме взаимодействия в мире элементарных частиц?
- Как связаны между собой радиус действия сил и масса частиц, обменом которыми обусловлено данное взаимодействие?
- Что такое комптоновская длина волны частицы? Как она связана с пространственным масштабом взаимодействия, осуществляемого обменом такими частицами?
- Какую роль играет малый радиус слабого взаимодействия в эволюции Солнечной системы?
- Почему единая теория электрослабого взаимодействия актуальна только при высоких энергиях? При каких именно?
- Почему протоны и нейтроны не являются в истинном смысле элементарными частицами?
- Какую роль играют соображения симметрии в теории элементарных частиц?
- Покажите, что невозможна аннигиляция электрона и позитрона с рождением только одного кванта.